STATISTICS

a unit for INTRODUCTORY PSYCHOLOGY

A Behavioral Research Laboratories Program / **ronald a. kinchla**

AN ADDISON-WESLEY PROGRAMMED TEXT

A PROGRAMMED TEXT

STATISTICS

a unit for INTRODUCTORY PSYCHOLOGY

RONALD A. KINCHLA
Assistant Professor
Department of Psychology · New York University

Under the auspices of
BEHAVIORAL RESEARCH LABORATORIES
PALO ALTO, CALIFORNIA

ADDISON-WESLEY PUBLISHING COMPANY, INC.
Reading, Massachusetts — Palo Alto — London

Palo Alto, California

Printed in the United States of America

The author expresses his deep gratitude to the staff of Behavioral Research Laboratories for its cooperation during the long months in which this manuscript was being prepared. I am particularly grateful to Dr. Allen D. Calvin for his guidance and assistance with the writing of the program. Finally, I should like to thank Miss Lynn Overton and Miss Mary Campbell for their help in the preparation of the manuscript, as well as the many students and teachers who participated in the testing of the program.

Ronald A. Kinchla

TABLE OF CONTENTS

TABLE OF CONTENTS

TO THE INSTRUCTOR

This program introduces the student to the elements of **statistical reasoning** and the manner in which this form of reasoning enters into the process of **behavioral research**. The material is presented in a carefully pre-tested sequence. Fundamental topics such as **variables**, **values**, and **distributions** are broken down into a series of small, sequentially organized steps. The student is led into discussion of such relatively complex concepts as **decision rules** and the probability of **Type One** and **Type Two errors** only after the prerequisites for his comprehension have been developed in detail.

After dealing with data as a collection of observed values of a variable, distributions of values are considered along with the manner in which **descriptive statistics** characterize various features of these distributions. **Formulas** for the **mean** and **variance** are developed in detail, and in a way that clearly indicates the exact feature of a distribution represented by each expression.

The concept of a **sampling distribution** is the central theme of the sections on **statistical inference**. Thus, the effects of sampling procedure and sample size in determining distributions are discussed in detail. In addition, the role of **probability theory** in the calculation of theoretical sampling distributions is considered as a basis for using random sampling procedures.

In programmed learning, each student proceeds at his own rate. In addition, the material is constructed so that **the student actively participates in the learning process**, supplying answers which require his understanding of that item. The answers constitute immediate feedback and reinforce the student's correct responses.

Periodic tests covering the material are provided to allow the student to evaluate his understanding as he progresses. Two forms of these exams are included in the instructor's manual.

TO THE STUDENT

Your programmed text is quite different from an ordinary book. A program consists of a large number of "frames," or numbered statements, each of which tells you something and asks questions about the material you have learned. The frames introduce new material a little at a time and review old material as needed to make sure you will remember it.

You do not study a program in the same way you study a book. The program is designed to let you work at your own rate of speed.

Get ready to use your program by covering the answer column at the right with the slider provided for you. Next, read the first frame and write the answer either in the blank or on a separate piece of paper. Then, move your slider down to uncover the answer and see if you are right. Go on in the same way to the following frames, checking your answer to each frame before going on to the next. Practice with the following frames.

1.	**When a blank has nothing under it, you simply fill in whatever best fits the blank. For example, the day of week that follows Thursday is ____________.**	**Friday**
2.	**When a blank has two answers underneath it, you select the one that best fits. For example, a dog ______ (is/is not) an animal.**	**is**

You will find that the left-hand pages of the program are upside-down and backwards. Pay no attention to them until you have finished all of the right-hand pages.

Turn to Frame 1 and begin. . . . . . .

Section I: **Data**

1. Suppose you were a psychologist studying how accurately a person judges distance. You might place a target some distance in front of a subject and ask him how far away it was. You would be interested in the difference between the **distance** he reported and the true ______________ of the target. distance

2. You might decide to move the target after each of the subject's judgements, asking him to judge each new distance. In your experiment, you would be collecting many different judgments of distance by changing or varying the ______________ between the subject and the target. distance

3. We refer to something that does not change during an experiment as a **constant.** Since the distance between the subject and the target is varied or changed, it ______________ (would/would not) be a constant in the experiment we just described. would not

4. We refer to the opposite of a constant as a **variable.** Something that changes in an experiment is called a variable. In the experiment we just described, the distance between the subject and the target is varied; therefore, we would refer to the varying distance as a ______________ (constant/variable). variable

7. Another name for a standard score from a normal distribution is:

a. an A score.

b. a Z score.

c. a relative score.

d. none of the above ______ b

8. A normal distribution is often described as being shaped like a:

a. square.

b. circle.

c. triangle.

d. none of the above ______ d

TRUE OR FALSE

9. The formula for s^2 is $\frac{\sum(X - \bar{x})^2}{N}$ ______ false

10. Rejecting an hypotheses which is really true is referred to as a Type 1 error. ______ true

11. Failing to reject the hypothesis which is really false is called a Type 2 error. ______ true

5. The target distance would be called a variable in your experiment because this distance was __________ during the experiment. changed

6. As a psychologist, you might wish to study how accurately a person judges weights. You might give the subject a lead ball and ask him to fill up a bag with sand until the bag felt as heavy as the lead ball. When he was finished trying to match the weight of the sand bag and the lead ball, you could compare the _________ of the bag of sand and the weight of the lead ball. weight

7. You could repeat this procedure several times, each time changing or varying the size of the lead ball. In this experiment, the weight of the lead ball would be referred to as a _______________. variable
constant/variable

8. The weight of the lead ball was a variable in this experiment since this weight was __________________ during the experiment. changed

9. Suppose you conducted the experiment **differently**. Suppose you gave the subject the **same** lead ball each time. You would not expect the subject to produce a bag of sand weighing exactly the same amount each time, even though the ______________ of the lead ball remained the same throughout the experiment. weight

If you weighed each bag of sand carefully, you would doubtlessly find that the weight differed slightly for each bag. Therefore, in this experiment, the weight of the lead ball was a _______________, whereas the constant
constant/variable
weight of the bag of sand the subject produced each time was a _______________. variable
constant/variable

REVIEW IV

FILL IN THE BLANKS:

1. A process in which there is uncertainty concerning outcomes can often be described as a ________ process. random

2. The two outcomes from the toss of a coin, heads or tails, ________ (can/cannot) be said to be mutually exclusive. can

3. All of the probabilities in a probability distribution will add up to ________. 1

4. The largest number you could assign to any member of a sample space would be ________. 1

MULTIPLE CHOICE:

5. In order for a list of outcomes to be a list for a sample space, it must be:

   a. inclusive and complete.

   b. exhaustive and mutually exclusive.

   c. inclusive and exhaustive.

   d. none of the above ________ b

6. The probability of rolling a die and obtaining more than one spot would be computed by using the:

   a. addition rule.

   b. multiplication rule.

   c. subtraction rule.

   d. division rule. ________ a

10. As a psychologist, you might be interested in how well people can remember things. For example, you might read a list of six letters of the alphabet to a subject, such as **m**, **t**, **s**, **p**, **g**, **k**, and then ask him to repeat this list. Suppose the subject said that the letters were **m**, **t**, **s**, **p**, **g**, **d**. He would have repeated only the first _____ letters correctly. 5
4/5

11. Suppose you then read another list of 6 letters — for example **p**, **f**, **t**, **m**, **s**, **r**. If subject's response was **p**, **f**, **t**, **g**, **f**, **n**, he would have repeated only the first _____ letters correctly. 3

12. You could conduct an experiment in which you gave the subject many different lists of **six** letters, each time asking him to repeat as many of the letters as he could. While the actual list of letters would be varied in this experiment, the number of letters in each list would always be six.

Since the number of letters in each list would always be six, the number of letters would be a _______________ constant
constant/variable
in this experiment.

13. Every time you give the subject a list of letters, you could record how many letters he repeated correctly before making an error. You would call this his **score** on each list. Therefore, if he repeated four of the six letters correctly before making an error, his score for that list would be ______. 4

14. If you read the subject the letters **m**, **p**, **t**, **x**, **z**, **s**, and his response was **m**, **p**, **t**, **x**, **z**, **d**, his score on that list would be ______. 5

245. This simply means the **decision rule** you used to reject the hypothesis was such that if the experiment were repeated a large number of times and the hypothesis were actually true, you would erroneously reject it about ______ times out of every 100 replications of the experiment. 5

246. This concludes our program. It is the author's hope that the information you have obtained will not only increase your understanding of Statistics, but your appreciation of its power as a tool in behavioral science.

15. Since every list contains 6 letters, whenever the subject repeated the list without making an error, his score would be _____. The worst possible score he could make would be zero, which would mean he had repeated _______ (none/all) of the letters correctly. — 6 — none

16. You know, therefore, that all of the subject's scores will will be no **greater** than _____ and no **less** than ___. — 6, 0

17. Because the subject's score can vary or change from list to list, it would be called a ____________ in this experiment. — variable

18. Considering his score as a variable, list below the possible scores that the subject could make (beginning with the worst score and moving in order to the best).
____ — 0, 1, 2, 3, 4, 5, 6

19. We will refer to each of these possible scores from 0 to 6 as a **possible value** of this variable. Thus, the smallest possible **value** of this variable is _____, and the largest possible **value** is _____. — 0 — 6

The subject _____________ (could/could not) receive a score of 10 because the list contains only six letters. Therefore, 10 _________ (is/is not) a possible value of this variable. — could not — is not

20. We have used the word _______________ to refer to things that may change or vary during an experiment. We use the word __________, however, to refer to things that do not change during an experiment. — variable — constant

240. The important features of estimation were **the facts** that you not only wished to choose the **best** estimate of the population statistic from your sample, but (if you knew something about the sampling distribution of that sample statistic) you could specify something about the a______________ of your estimate.

accuracy

241. You specified the **accuracy** of the estimate by noting the probability of obtaining any particular size difference between your estimate and the ____________________ ______________.

population statistic

242. In order to test an **hypothesis**, you first stated a ______________ rule which indicated how __________ an observation would have to be in order to convince you that the hypothesis should be rejected.

decision, unusual

243. If you could determine a theoretical sampling distribution which indicated the probability of each possible observation under the assumption that the hypothesis was true, you could specify the probability of erroneously **rejecting** the hypothesis on the basis of your decision rule. In other words, this theoretical sampling distribution allowed you to specify the risk of making a Type __________ error.
One/Two

One

244. If you **reject** an hypothesis on the basis of the outcome of an experiment, and the **decision rule** you used involved a risk of making a Type One error of .05, you can say the results of the experiment were s__________ at the _______ level.

significant
.05

21. One of the variables we considered was distance. A **value** of that variable would be any particular distance. Since 10 feet is a particular distance, it would be a ________________ of that variable. value

22. Another variable we considered was weight. Values of that variable could be particular ____________ — such as 10 pounds, 2 ounces, 3 pounds, etc. weights

23. There are many different things which could be variables in an experiment. All variables, however, have one thing in common. They are things that may ________________ during an experiment. change

24. Since many things may be variables, it is often useful to give each variable a **name**. For example, "distance" is the name we used for one variable we considered, whereas "weight" is the n____________ of another variable we considered. name

25. Any particular **distance** is a value of the variable named "distance." Any p______________ weight is a value of the variable named "weight." particular

26. Two feet ______________ be a value of the variable named weight. would not
(would/would not)

27. In the experiment in which the subject tried to repeat a list of six letters, we considered a variable named "score." The subject's "score" was the number of letters he repeated correctly before making an error. The different possible __________ of this variable were 0, 1, 2, 3, 4, 5, 6. values

236. In reference to our illustration, suppose the difference in the mean test scores between the group of specially trained and the group of normal students was so **unusual** that it would occur less than 5 times out of 100. If the hypothesis really were true (if both student groups were really samples from the same population), you might reject the hypothesis that there was **no effect** from the training and say the results were s____________ at better than the .05 level. — significant

237. Whenever you **reject** an hypothesis on the basis of a decision rule whereby the risk of a Type One error is .05, you can say that the results of the experiment were **significant** at the .05 level. This simply means the **decision rule** you used to **reject** the hypothesis would cause you to erroneously reject it about _____ times out of 100. In other words, the risk of a Type One error was _____. — 5 — .05

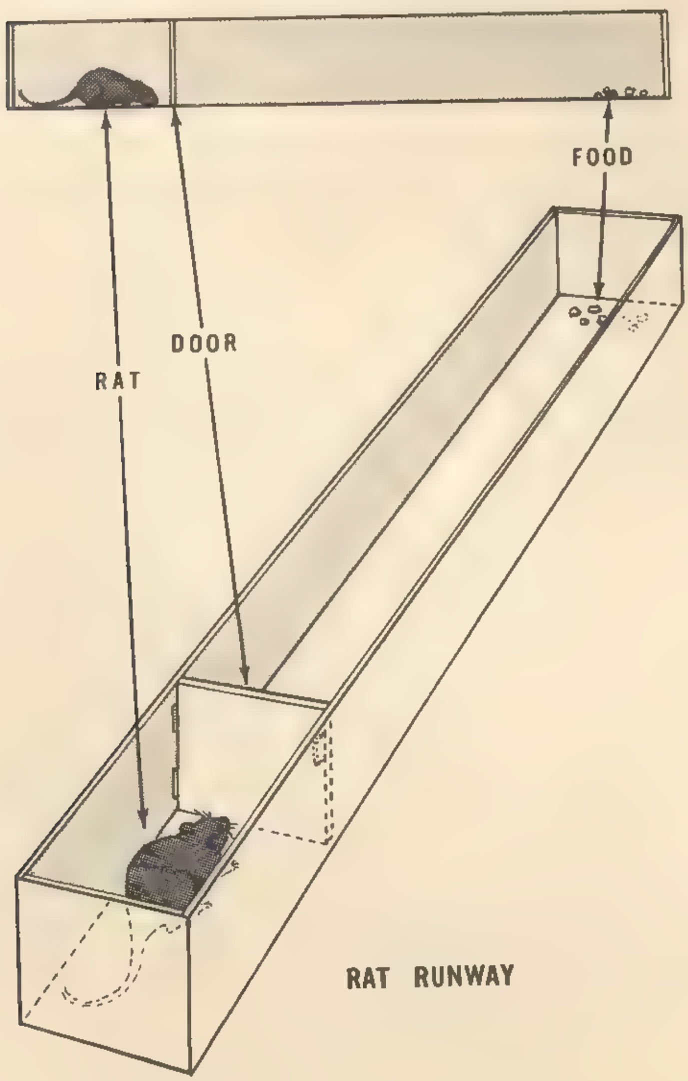
238. If you **rejected** an hypothesis on the basis of a **decision rule** whereby the risk of a Type One error was .01, you would say the results of the experiment were s____________ at the ________ level. — significant, .01

239. We have considered two types of statistical inference in this section. One was the problem of ____________ population statistics on the basis of a sample. The other was the procedure for testing an ____________ concerning a population on the basis of a sample. — estimating — hypothesis

28. Therefore, if the subject's score changed from 3 to 5, we would say that the __________ of the variable had changed. value
value/name

29. You have learned that a ______________ is something that can change during an experiment. Every time a variable changes, it changes from one _________ to another. variable value

30. Psychologists have studied how fast a rat will run down a narrow alley to secure food. The picture below shows an experimental setup you might have used if you were studying running speeds of rats.



RAT RUNWAY

232. However, if they were randomly drawn from the same population, you ______ (would/would not) expect to find a very large difference between the two means. would not

233. If the preceding hypothesis were true (the test performance of each group of students could be viewed as a sample from the same population), you ______ (would/would not) expect one group of students to do a great deal better than the other group. would not

234. Since the distribution of **differences** between the means of pairs of samples drawn from the same population **can** be derived by probability theory, it is possible to determine how **unusual** any particular size difference would be if the hypothesis were true. There is general agreement among psychologists as to **how** unusual a particular difference of this sort would have to be in order to reject the hypothesis. A decision rule is often used whereby the hypothesis is **rejected** if an observation is made which would occur fewer than 5 times out of 100 were the hypothesis true. In other words, if the outcome of the experiment has a probability of ______ or less were the hypothesis true, the hypothesis would be rejected. .05

235. A research worker often says a result is **significant** at the .05 level if the result of the experiment has a probability of .05 or less when the hypothesis is true. Therefore, saying a result is **significant** at the .01 level implies that the result of the experiment would have a probability of ______ or less if the hypothesis being tested were true. .01

30. (continued)

The rat is placed in one end of the runway and food is placed in the ___________ end of the runway. other
same/other

When the door is opened in front of the rat, he is free to move all the way to the far end of the runway to secure the ______________. food

You would be interested in the time between the opening of the ______________ and the time when the animal reached the food. door

31. Let's imagine that you are conducting the following experiment. Suppose you took a rat that normally ate four ounces of food a day and that had never been in the runway before. You begin to feed the rat only in the runway. At precisely the same time each day, you place the rat at the starting point of the runway, four ounces of food at the other end. Then, you open the door of the runway. Each day, you would record the ______________ between the opening of the door and the rat's reaching the food. time

32. In this experiment, you would be interested in a ______________ named **running time.** variable
variable/constant

33. The amount of food placed in the runway each day would be a ______________ in this experiment. constant
variable/constant

229. If it is possible to determine the theoretical sampling distribution when the hypothesis is true, we can determine the probability of a Type ________ error (erroneously **rejecting** an hypothesis) for any particular d__________ rule we might decide to use.

One
decision

230. This is another example of a procedure for making a statistical ________________ about a population on the basis of a sample.

inference

231. This particular type of **statistical inference** is very useful in evaluating the results of experiments. Very often, the hypothesis being tested is whether or not the experiment demonstrates a real difference between two experimental conditions. For example, a special reading course could be given to a group of students and the results on a final reading exam could be compared with the results of a regular reading class. You could state the following hypothesis to be tested.

> **Hypothesis**: There was no advantage in taking the special reading course (therefore, the performance of the 2 groups of students can be viewed as 2 samples from the **same** population).

An hypothesis of this sort can be shown to indicate a particular theoretical sampling distribution that is approximately normal and that is called the sampling distribution for the **difference** between two sample means drawn from the **same** population. Even if two random samples were drawn from the same population, you ________________ expect them to have exactly the
would / would not
same mean.

would not

34. Suppose you repeated this experiment for 10 days. The record of your results might look like this:

DAY	RUNNING TIME
1	200 sec.
2	100 sec.
3	150 sec.
4	80 sec.
5	40 sec.
6	41 sec.
7	15 sec.
8	10 sec.
9	4 sec.
10	3 sec.

On day 1, the rat took 200 sec. from the time the door opened to reach the food. On day 2, he took 100 sec. On day 3, he took ________ sec. 150

35. The rat's running time on day 10 was ___________. 3 sec.

36. The rat's running time was _____________ on day 7 than on day 3. shorter
longer/shorter

36. Running time was _____________ on day 3 than on day 2. longer
longer/shorter

37. Instead of referring to "the variable named running time," we shall simply say "running-time" variable. Thus, the variable we call "running time" will be referred to as the "___________ - ______________" variable. running-time

38. **Particular** running times are _____________ of the "running-time variable". values

225. According to the decision rules represented by Graph B, the rule should read as follows:

Decision Rule: If the subject makes more than _____ correct responses, reject the hypothesis that he is only guessing; otherwise __________ it. — 5 — accept

226. Notice the decision rule represented by Graph A would lead to a ______________ (greater/smaller) risk of erroneously rejecting the hypothesis than would the decision rule represented by Graph B. — smaller

227. The most important feature of the previous illustration was the fact that the hypothesis being tested actually indicated a particular **theoretical sampling distribution**. This made it possible to evaluate the probability of a Type One error for particular ______________ rules. — decision

228. Let's review the major steps in testing an hypothesis on the basis of an observation. Basically, we are trying to determine if a particular hypothesis is consistent with what we have observed. If the observation would have been highly unusual were the hypothesis true, we would decide to ____________ (accept/reject) the hypothesis. — reject

On the other hand, if the observations would not have been unusual were the hypothesis true, we probably would ____________ (accept/reject) the hypothesis. — reject

39. A large part of any scientist's work consists of observing the variables in experiments and making records of these observations.

Psychologists are chiefly interested in **behavior** — either the behavior of human beings or the behavior of animals. Therefore, if you were a psychologist, a large part of your work would probably consist of observing b____________ and making records of these observations. — behavior

40. You might be interested in observing how quickly a person could solve a mathematical problem, how accurately he could estimate the weight of some object, or how much time he spent sleeping. Whatever behavior you were interested in, you would probably make ob____________ of this behavior and records of your ob____________. — observations; observations

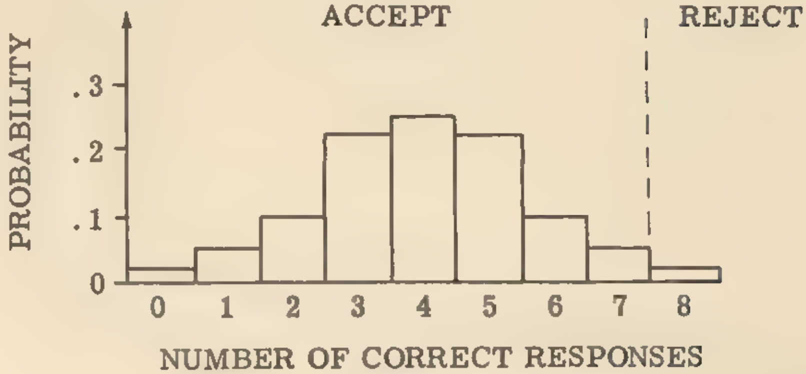
41. Each of the variables we have discussed has had different **possible** values. For example, if you tossed a coin in the air and it fell on one side or the other, the two possible results you could observe would be a "head" facing up or a "__________" facing up. These two possible outcomes are the possible values of the variable we would call "results of a coin flip," or perhaps "**falls of a coin.**" — tail

42. What if the coin were tossed three times and you **observed** that it fell with the "head" facing upwards all three times? The **observed values** of the variable "falls of a coin" would be "heads," "heads," and "heads." Even though you did not actually observe a "tail," the two ____________ values of the variable are still "heads" or "tails." — possible

possible/observed

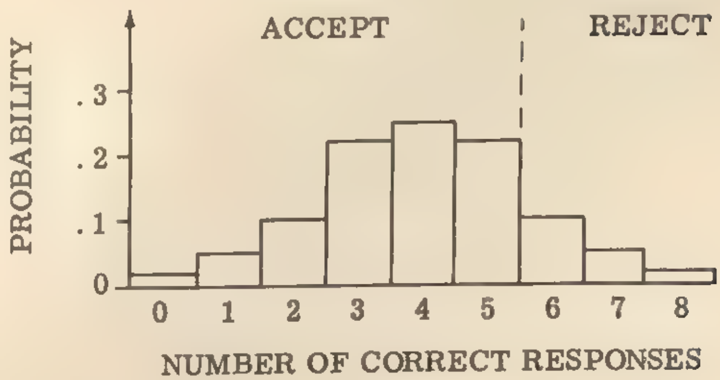
224. The following two graphs indicate the same probability distribution described in the previous table. We have indicated, however, a different decision rule on each graph.

Theoretical sampling distribution for the discrimination experiment assuming pure guessing



GRAPH A

Theoretical sampling distribution for the discrimination experiment assuming pure guessing



GRAPH B

The following decision rules is represented by Graph

__________.
A / B

A

Decision Rule: If the subject makes more than 7 correct responses, reject the hypothesis that he is guessing; otherwise, accept it.

43. Suppose you were conducting a telephone survey in which you asked each person you called whether or not he had been watching television. Assuming each person answered the question, the two **possible** answers are "yes" or "no." Thus, the two **possible** values of the "answer" variable are "________" and "________." yes, no

44. Suppose four people were called in this telephone survey. Suppose you **observed** that the first person said "yes," the second person said "no," the third person said "yes," and the fourth said "yes." The record of your observations would be:

	Response
1st person	yes
2nd person	no
3rd person	yes
4th person	yes

These four answers are the ________________ (observed/possible) values of the variable under study. observed

45. Imagine that you were learning how to bowl and that you decided to keep track of how you improved with each lesson. As a test of your skill, you could bowl 5 balls following each lesson. After each ball, you could reset any pins you had knocked down so that exactly 10 pins were standing each of the five times you rolled a ball. The most pins you could knock down with any one ball would be _______ pins. The worst you could do with any one ball would be to knock down _______ pins. 10 0

220. We are now in a position to define a **decision rule** and to evaluate the resulting probability of erroneously rejecting the hypothesis that the subject is simply guessing. For example, consider the following __________ ______. decision rule

> If the subject makes more than 6 correct responses, reject the hypothesis that he is only guessing; otherwise, accept the hypothesis.

221. According to the previous table, the probability that a subject who is just guessing will obtain 7 correct responses is __________ and the probability of his obtaining 8 correct responses is ________. .031 .004

222. Therefore, the probability of erroneously rejecting the hypothesis that he is only guessing (the probability of a Type _________ (One/ Two) error) is equal to the probability of a subject who is only guessing obtaining either 7 **or** 8 correct responses. According to the addition rule of probability, this is equal to _______ + _______, or ____. One 031, .004, .035

223. It is a characteristic of normal human hearing that there are some tones a subject can hear sometimes and not at other times. In other words, it is possible for him to hear the tone on some trials and yet not hear the tone on other trials. Thus, a subject who made fewer than 8 correct responses ______________ (could/ could not) have heard the tone on some trials and not be **purely guessing**. could

46. If we consider the number of pins you knock down with each ball to be a **variable**, the smallest possible value of the variable would be _______ and the largest possible _____________ of the variable would be 10. — 0, value

47. List all the possible values of the "pins-knocked-down" variable, starting with the smallest possible value and ending with the largest possible value.

___, ___, ___, ___, ___, ___, ___, ___, ___, ___, ___ — 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

48. Suppose after the first lesson you rolled 5 balls with the following results:

PINS KNOCKED DOWN	
1st Ball	0
2nd Ball	0
3rd Ball	2
4th Ball	10
5th Ball	5

According to this list of results, you knocked none of the pins down with the first ball, none down with the second ball, _________ down with the 3rd ball, ______ down with the 4th, and _______ down with the 5th ball. — 2, 10, 5

Thus, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 are the _________________ (possible/observed) values of the variable we are studying (the "pins-knocked-down" variable), and 0, 0, 2, 10, and 5 are the _________________ (possible/observed) values of that variable. — possible, observed

215. If the subject were just guessing, the most probable number of correct responses would be ______, since it has a probability of ______. 4 .273

216. The probability of a subject obtaining 4 correct responses when he is only guessing is .273. This means that **about** ________ out of every 1,000 subjects who were purely guessing would be able to make 4 correct responses in 8 trials. 273

217. While subjects could quite commonly make 4 correct responses if they were just guessing, it would be __________ (usual/unusual) for a subject who is just guessing to make as many as 8 correct responses, since the probability of obtaining 8 correct by guessing is ________. unusual .004

218. You would expect only about _____ out of every 100 students who were just guessing to get as many as eight correct, since the probability of getting 8 out of 8 correct (if you are only guessing) is .004. 4

219. If a subject obtained an unusually **small** number of correct responses, it would seem reasonable to suppose that he was just unlucky and really couldn't hear the tone. Otherwise, you would have to assume he could hear the tone but had decided to make just the opposite response that he should have made. Therefore, we shall assume that an unusually **small** number of correct responses indicates that the subject ____________ (could/could not) hear the tone. could not

49. Each **observed** value of a variable is referred to as a single **observation.** Therefore, in the illustration we just considered, the five observed values of the "pins-knocked-down" variable would be referred to as five ______________. observations

50. If you were studying how fast a rat ran down an alley to secure food and observed that it took him exactly 10 seconds, this time would be an observed value of the "running-time" variable and ______________ (would/would not) be referred to as a single observation. would

51. It is possible the rat could take as long as 20 minutes to reach the food. Suppose, however, that the longest observation of running time was 3 minutes. Therefore, whereas _____ (20/3) minutes would be a possible value of the running-time variable, _____ (20/3) minutes would be both a possible value and an observed value. 20 3

52. When a particular variable is observed or studied in an experiment, records are made of the observed values of this variable. These records are called **data.** For example, we just considered an experiment in which a rat ran down an alley to get food. The time it took for the rat to reach the food each day was observed and recorded. Our records of these running times are the d________ from that experiment. data

53. Another experiment we discussed dealt with a subject's repeatedly attempting to match the weight of a lead ball by filling a bag with sand. Records of the weight of each bag of sand the subject produced are the ________ from that experiment. data

212. If we assume that the subject was simply guessing, the probability of a correct response on each trial would be one-half (just as if he were trying to predict how the fair coin would fall). The following table indicates the theoretical sampling distribution indicated by this hypothesis.

Theoretical sampling distribution for the discrimination experiment assuming pure guessing (no discrimination).

Number of Correct Responses	Probability
0	.004
1	.031
2	.109
3	.219
4	.273
5	.219
6	.109
7	.031
8	.004

This is a binomial distribution indicating the probability of each possible number of successes in 8 trials where the probability of a success on each trial is one-half.

213. The nine possible outcomes of our experiment indicated in the first column of the table are based on the ________ of correct responses the subject made in the 8 trials. — number

214. You can think of these nine possible outcomes of the experiment as a **sample space**. Thus, the numbers shown in the second column of the table are a p________ d________ on the sample space consisting of the 9 possible outcomes of the experiment. — probability distribution

54. Notice that we referred to the ______________ values
possible/observed
of the variable as the data. A list of the **possible** values of a variable would not be considered data. | observed

55. Suppose you tossed a coin two times and the observed value of the variable "falls of the coin" was "heads" on both tosses. Then "heads" and "tails" are the two ______________ values of the variable, whereas | possible
observed/possible
"heads" and "heads" are the two ______________ values. | observed
possible/observed

56. In the preceding frame, the list of values we would call **data** was ______________. | "heads" and "heads"
"heads" and "heads"/"heads" and "tails"

57. "Heads" and "heads" would be considered **data** since they are the ______________ values of the variable. | observed
possible/observed

58. Television stations are naturally interested in which programs are preferred by television viewers. Imagine that you were hired to determine viewing preferences in an area in which there were only three television channels: Channel 5, Channel 7, and Channel 9. Suppose you asked a number of television viewers the following question: "If you had to watch only one of the three television stations for a week, which one would you select?" Acceptable answers to these questions would be Channel _____ or _____ or ___. | 5, 7, 9

59. If you considered the viewer's answer as a variable, the three possible values of the variable would be 5, 7, and _____. | 9

208. If a subject were actually deaf to a tone of this frequency and based his responses on pure guessing, rather than on what he actually heard, he would have the same probability of being correct on each trial as he would if he were ________________ the outcome of tossing a ______________. — guessing / coin

209. Assuming the experimenter's coin was a fair one, you know the probability of **guessing** how it would fall is exactly ________________, since you would expect to be correct just about as often as you would expect to be wrong. — one-half

210. If you assumed the subject's performance on these 8 trials was the product of pure guessing, it would be possible to calculate a theoretical sampling distribution for the number of correct responses out of 8 trials, since it would be possible to choose an appropriate b________ d________________ to represent the probability of each number of correct responses in N trials. — binomial / distribution

211. The following hypothesis actually indicates a particular theoretical sampling distribution for each possible number of correct responses in the 8 trials of the experiment.

Hypothesis: The subject's responses are based on pure guessing (the probability of a correct response on each trial of the experiment is __________). — one-half

60. While you probably would ask many people this question in a real study, let us suppose you asked only three people and the first person selected 7, the second person selected 9, and the third person selected 9. The values 5, 7, and 9 would be the ________________ values of the "answer" variable, whereas the values 7, 9, and 9 would be the ________________ values of the variable. — possible

(possible/observed)

observed

Since an observation is any observed value of a variable, the three observations are _____, _____, and 9. — 7, 9

Your **data** would be the values _____, _____, and _____, since these were the ________________ values of the "answer" variable. — 7, 9, 9 — observed

(observed/possible)

61. In the television survey, the answer of each of the three people was a ________________, but the question they were asked was a ________________. — variable — constant

(constant/variable)

(constant/variable)

62. It is often useful to distinguish between **continuous** and **discrete** variables. If you were to count the number of people in a room, it would be possible to have 8 people or 9 people, but it would not be possible to have $8\frac{1}{2}$ people. It is this characteristic of the variable named "number of people in a room" which makes it a **discrete** variable. The variable is discrete since there are no values of the variable between 8 and 9, or between 10 and 11, or between 12 and ______. — 13

205. (Continued)

cycle per second tone (a very high-pitched tone which very few people can hear), you decided to conduct the following experiment. Each student would be a subject in a brief experiment consisting of 8 trials. On each of these **trials** he would be asked to determine whether or not you had presented a brief burst of the 20, 000 cycle per second tone. He was required to say "yes" or "no" to indicate whether or not he thought the tone had been presented. In other words, he would say "_____" if he thought the tone had been presented and "_____" if he thought it hadn't (on each of the 8 t_______).

yes
no
trials

206. The data from a particular subject would consist of ____ observations of a variable called "his answer," whose two possible values are "________"-or "______."

8
yes, no

207. To discourage the subject from simply guessing whether or not you had presented the tone on each trial, you informed all of the students that you would determine whether or not to present the tone on each trial by secretly tossing a coin at the beginning of each trial. If the coin landed "heads," you would present the tone; and if it landed "tails," you would not. In other words, if a student could **not** hear a tone of such a high pitch and attempted to **guess** whether you actually presented a tone, he would be in the same position as if he tried to _____________ how a coin would land when it was tossed.

guess

62. (Continued)

On the other hand, the variable we call "length" is an example of a **continuous** variable. No matter which particular pair of lengths you considered, it

______________ be possible to imagine a length **between** — would
would/would not

these two lengths. For example, $8\frac{1}{2}$ inches is between 8 inches and ________ inches. (By "between", we mean — 9
7/9

greater than one length but less than the other.)

Similarly, $2\frac{1}{2}$ feet is a value of the variable called "length," which is b__________ the values 2 feet and 3 feet. In other words, $2\frac{1}{2}$ feet is larger than 2 feet but less than 3 feet. — between

63. A **continuous** variable has an unlimited number of values because no matter how close two values are to each other it ________ always possible to imagine another — is
is/is not

value which would lie between them. For example, even though 2.10 inches and 2.20 inches are close together, __________ inches is between them. On the — 2.15
2.15/2.25

other hand, if the variable you are studying is **discrete**, you can find two values of the variable such that there is no value between them. For example, the number of pennies you have in your pocket would be a value of a ________________ variable, since you could only — discrete
discrete/continuous

have in your pocket one penny, two pennies, three pennies, and so on. Of course, you could not have any number **between** these values.

202. On the other hand, while **decision rule B** reduces the risk of making a Type Two error (erroneously **failing** to **reject** the hypothesis), it does produce the largest _____ of making a Type One error (erroneously **rejecting** the hypothesis), since it will result in ___________ (more/fewer) non-defective parachutes being shipped back to the factory than will the other rule.

risk

more

203. The preceding illustration once again indicates how the particular d_______________ r_________ used in **hypothesis testing** determines the risk of making the two possible kinds of errors.

decision rule

204. The most important contribution of a statistical approach of this kind to decision-making is your ability to **specify** the relation between particular _________________ _____________ and the risk of making particular kinds of errors.

decision rules

205. In the preceding illustration we obtained the theoretical sampling distribution by estimating it from an experimental sampling distribution. It is possible to test certain **hypotheses** for which a theoretical sampling distribution is actually defined. In fact, it is highly desirable to phrase the question about an experimental observation in such a manner that the hypothesis to be tested actually specifies a theoretical sampling distribution. The following illustrations will serve to demonstrate the manner in which an hypothesis may actually define a particular theoretical sampling distribution. Suppose you were a psychologist who was interested in the ability of people to hear very high-pitched tones. In order to determine how many people in your psychology class could actually hear a 20, 000

64. No matter how similar the weights of two objects, it would always be possible to imagine an object whose weight was less than one but greater than the other. Therefore, "weight" is an example of a ________________ variable. continuous
discrete/continuous

65. It will not be necessary to distinguish between discrete and continuous variables very often. Most of the illustrations in this text involve discrete variables. Those illustrations that involve continuous variables are treated as if they were discrete variables. For example, if you measured a person's height to the nearest inch, you might say he was 65 inches tall or 66 inches tall, because you "round off" his "height" to the nearest inch, you would never say he is 65 $\frac{1}{2}$ inches tall. In other words, you are treating the
________________ variable called "height" as if it continuous
continuous/discrete
were a ________________ variable. In other words, discrete
discrete/continuous
you are pretending that there are no values between 65 inches and 66 inches, or between 67 inches and 68 inches, and so on.

66. To determine the **number** of people in a room, you would count them. To determine the number of eggs in a basket, you would count them. Whenever you ____________ things, you are determining how many (what number) of things there are. count

67. Any collection or group of things can be counted. The procedure we call **counting** tells you the **number** of things that are in the group or collection. For example, we can determine the ______________ of windows in a building by counting them. number

198. (Continued)

Once again, the shaded columns in each graph represent the probabilities of ____________ rejecting the
correctly/ erroneously
hypothesis because you obtained a non-defective parachute with an unusual weight. Since the shaded columns in each graph represent the probability of making a Type One error, you could add these probabilities according to the simple addition rule to find the probability of making a Type _____ error with each decision rule.

erroneously

One

199. The preceding graphs should make apparent the fact that the probability of erroneously rejecting the hypothesis (the probability of a Type One error) is ____________
smaller/ larger
if you use decision rule A than if you use decision rule B.

smaller

200. The risk of a Type One error (erroneously rejecting the hypothesis) is greatest under decision rule B, since even slightly unusual weights will result in a rejection of the hypothesis. In fact, only parachutes weighing 44, 45, or 46 ounces would lead you to ____________
accept/ reject
the hypothesis and release the parachute for use by an airman.

accept

201. Because you are much more concerned with avoiding Type Two than you are Type One errors, you would probably use decision rule _____, since it would result
A/B
in the smallest risk of making a Type Two error.

B

68. It is often useful to distinguish between the name of something and the thing itself. A "name" is something you speak, write, or read. For example, "Boston" is the name of a city. You could go to Boston, walk through the streets of Boston, or even live in Boston. On the other hand, the name "Boston" is something you speak, write, or read. Similarly, if you had a dog named "Rover," you might scratch the ________ (dog/name) behind the ears, whereas you might print his ________ (dog/name) on his dog house.

dog

name

69. Suppose a mother didn't decide to name her new baby "Kendall" until three days after the baby was born. The ________ (baby/name) would be three days old before it was given a name.

baby

70. We will use quote, " ", when we are referring to the name of something rather than to the thing itself. Thus, in the previous example we would refer to the baby as Kendall and to the name the baby received as "Kendall." Similarly, when we speak of the city of Boston and of the name "Boston," we will say that you could live in ________ (Boston/"Boston") and that you write the name ________ ("Boston"/Boston) as part of an address on an envelope.

Boston

"Boston"

71. The number of things in a group or collection is a characteristic of that group of things, just as the color of a person's hair is a characteristic of that person. We refer to a person's hair color with a **name**, such as "red," "brown," and "black." Similarly, we refer to the number of things in a group with a name, such as "five," "twenty," "eighteen," and "forty." Thus, "red" would be the **name** of a particular hair color and "five" would be the ________ of a particular number.

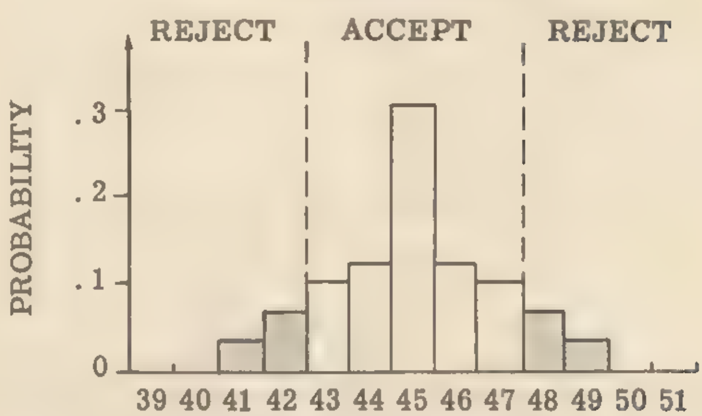
name

197. (Continued)

Notice that decision rule ______ (A / B) requires the weight of a parachute to be **more** than unusual before rejecting the hypothesis than does decision rule ______ (A/B). A B

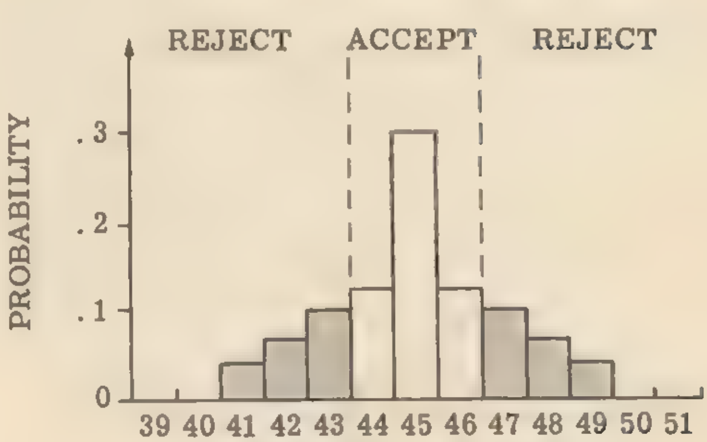
198. Decision rules A and B can be illustrated in the same manner as the previous decision rule by drawing vertical lines separating the possible weights into three classes, as shown below.

DECISION RULE A



WEIGHT IN OUNCES
(Rounded off to the nearest ounce)

DECISION RULE B



WEIGHT IN OUNCES
(Rounded off to the nearest ounce)

72. Names for other ______________ are "two," "6," "four," and "one hundred." numbers

73. There are different ways of representing the same number. For example, the ______________ of players on a basketball team could be represented either by the name "five" or by the name "5" or by the name "V." number

74. Number is a characteristic of a group or a collection of things. The names (such as "ten," "four," "6") that are used to represent this characteristic are often called **numerals**. Therefore, the number of players on a basketball team could be represented either by the numeral "5" or the Roman numeral _________. V

75. Since the names "red," "green," "Democrat," and "Boston" are not the names of numbers, they are not numerals. Therefore, of the two names "6" and "blue," "________" is a numeral since it is the name of a ____________. 6 number

76. We pointed out earlier that there is a difference between the name of something and the thing itself. A numeral is the name we give to a number. You determine the ______________ (number/numeral) of things in a group by counting them; you represent this characteristic of the group by writing or saying a ______________ (number/numeral). number numeral

77. The same number can be represented by different numerals. "Four," "4," and "IV" are three names for the same number; in other words, they are three **numerals** which represent the same ____________. number

195. We said earlier that a characteristic of this particular testing problem was the fact that we were much more concerned with avoiding Type Two errors than we were with avoiding Type One errors. In other words, we were **less** concerned with erroneous rejections of the hypothesis than we were with erroneous failures to reject the hypothesis. The reason for this was the fact that failing to reject the hypothesis when we should would result in an airman being issued a defective parachute, whereas ______________ the hypothesis when we shouldn't would only result in needlessly returning a parachute to the factory. (Remember, the hypothesis being tested is that the parachute is **not defective**.) — rejecting

196. You have already seen how it is possible to modify a decision rule so as to alter the risks involved in making the two possible kinds of error. Typically, if you decrease the risk of making one kind of error, you ________________ the risk of making the other type of error. — increase

197. Consider the two decision rules described below:

Hypothesis: The parachute is suitable (non-defective).

Decision Rule A: If a parachute weighs less than 43 ounces or more than 47 ounces **reject** the hypothesis; otherwise, accept it.

Decision Rule B: If the parachute weighs less than 44 ounces or more than 46 ounces, **reject** the hypothesis; otherwise accept it.

78. Number is often referred to as a **variable** — for example, the number of automobile accidents in California each year, the number of people who vote in an election, or the number of base hits a baseball player makes each game. In each of these cases, number could be thought of as a ______________. variable

79. Like any other variable, the variable we call number has different particular values. Just as "red" and "green" are **names** of particular values of the variable we call "color," — "three," "4," "20," and "ten" are n__________ of particular values of the variable we call number. names

80. Consider the following list of letters: **b**, **c**, **m**, **p**. We could determine the ______________ of letters in this list by **counting** how many there were. The name (numeral) we would use to represent this number would usually be "_____." number 4

81. "Four" and "4" are two names for the same ________. number

82. Numerals are simply names we use to represent different values of the variable we call __________. number

83. Roman numerals are names for particular numbers. Thus, the Roman numeral III represents the number we usually represent by the numeral "____". 3

84. Numerals are often used to represent characteristics of things other than number — such as weight, length, temperature, age, and so on. When we say that the temperature is "20 degrees below zero," we are using the ______________ (number/numeral) "20" as a name for a particular value of the variable we call temperature. numeral

190. The shaded columns in the previous graph represent the probability of a ____________ (defective/ non-defective) parachute having a weight which will cause us to reject the hypothesis that it is non-defective, since this is a distribution based on ____________ (defective/ non-defective) parachutes.

non-defective

non-defective

191. In other words, the shaded columns in the previous graph represent the probability of ____________ (correctly/ erroneously) rejecting a non-defective parachute because it has an unusual weight for a non-defective parachute.

erroneously

192. Erroneously rejecting an hypothesis when it is actually true is an example of a Type ______ (One/ Two) error.
(Remember, the **single** key word is rejecting.)

One

193. Therefore, the shaded columns in the previous graph represent the probability of a ______ ______ error.

Type One

194. Once again, we see how it is possible to evaluate a particular **decision rule** in terms of the probability of making a particular kind of error. On the other hand, since we **do not** know the distribution of weights among defective parachutes, we cannot specify the probability of correctly rejecting the hypothesis. In order to specify the probability of correctly rejecting the hypothesis (or of erroneously failing to reject), we would have to know the probability that a ____________ (defective/ non-defective) parachute would have a weight which would lead us to reject the hypothesis.

defective

85. When we use the names "twenty inches," "two feet," and "ten feet," we are using the ________________ numerals
numerals/ numbers
"twenty," "2," and "10" as **names** for particular values of the variable called length.

86. Football and baseball players often have
________________ written on the backs of their numerals
numerals/ numbers
uniforms to help people in the stands to identify the different players.

87. It is important to be careful when numerals are used to represent characteristics other than number. It makes sense to say that twenty things are **twice** as many things as ten things. It is also appropriate to say that something weighing twenty pounds is **twice** as heavy as something weighing ten pounds. It
____________ necessarily make sense to say that a does not
does/ does not
baseball player with the numeral "20" written on his jersey is **twice** as good as the baseball player with the numeral "10" on his jersey. The fact that one player was assigned the numeral "10" and another assigned the numeral "20" does **not** necessarily imply anything more about the players themselves than does the difference in their names (John and Charles, for example).

88. The value of a variable can often be represented by numerals. We shall refer to these variables as **numerical variables**. For example, we discussed an experiment earlier in which we were interested in the time it took a rat to run down an alley to reach food.

186. (Continued)

We have drawn vertical lines, dividing the possible weights into 3 groups. The group on the far left labeled "reject" represents all weights **less than** ______ ounces. 42

(We could describe this group of weights as consisting of all weights of "______ ounces or less" rather than saying "all weights less than 42 ounces.") 41

187. According to the previous **decision rule**, we would consider any parachute weighing 41 ounces or less to be so ____________ (usual/ unusual) that we would **reject** the hypothesis that the parachute was suitable (non-defective). unusual

188. Similarly, the group of weights at the far right of the graph consisting of all weights of 49 ounces or more would also be considered so unusual for a non-defective parachute that we would ____________ (accept/ reject) the hypothesis that the parachute was non-defective when it weighed 49 ounces or more. reject

189. On the other hand, any parachute weighing **more than** ______ or **less than** ______ ounces would not be considered sufficiently unusual (for a non-defective parachute) to cause us to reject the hypothesis. 41, 49

Therefore, this group of weights is identified by the word ____________ (accept/ reject), which is written between the vertical lines isolating this group of weights. accept

88. (Continued)

The values of this running-time variable were represented by numerals such as "20 seconds," "10 seconds," "800 seconds," Therefore, we _____________ refer to running time as a numerical
would/would not
variable. would

89. Variables whose values are not represented by numerals will be called **non-numerical variables**. If we were interested in hair color, therefore, the values we might observe would be black, red, brown, and so forth. Since these values are not represented by numerals, we will refer to hair color as a
_____________________variable. non-numerical
numerical/non-numerical

90. Political party is a variable, and particular values of this variable are Democrat, Republican, Socialist, and so on. This would be an example of a
_____________________variable. non-numerical
numerical/non-numerical

91. The age of American presidents when they were elected to office would normally be a _____________________ numerical
numerical/non-numerical
variable.

92. We said earlier that the **data** from a study was a record of the observed values of the variables being studied. If these observed values are represented by numerals, the variable under study is a _____________________ numerical
numerical/non-numerical
variable. We would refer to the data, therefore, as **numerical data**.

184. On the other hand, if a parachute weighed 45 ounces, you probably ______________ (would/would not) suspect that it was defective on the basis of its weight. would not

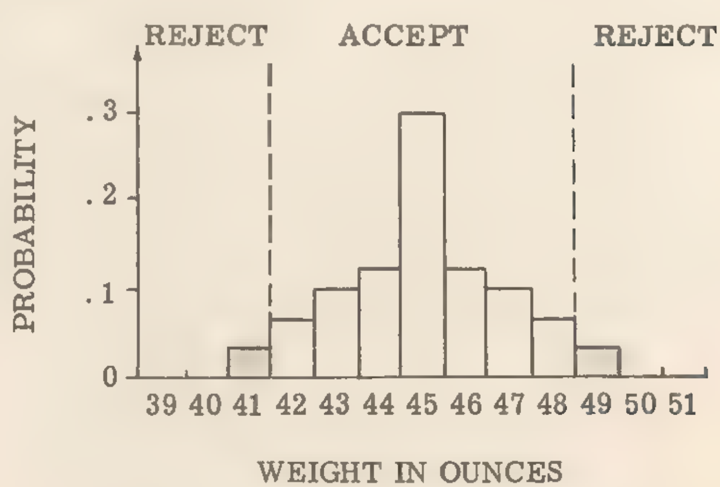
185. A parachute weighing 45 ounces would not be suspicious, since this was a typical or highly probable weight for a non-defective parachute. On the other hand, (considering the distribution of non-defective parachutes), if a parachute weighed 41 ounces or 49 ounces, its weight would be ______________ (typical/unusual) and you probably ______________ (would/would not) **suspect** that it was defective. unusual, wo

186. Imagine that you adopted the **decision rule** indicated below.

Hypothesis: Parachute is suitable (non-defective).

Decision Rule: Reject the hypothesis if the parachute weights **less than** 42 ounces or **more than** 48 ounces. Otherwise, accept the hypothesis.

The decision rule can be illustrated in terms of a previous distribution of weights for non-defective parachutes.



(Rounded off to the nearest ounce)

93. Records of the observed running times in the experiment in which we observed the time it took for a rat to reach some food would be

________________ data. — numerical
numerical/non-numerical

94. Earlier, we considered a study in which people were asked which of three television channels they preferred: 5, 7, or 9. We could name the three possible answers they gave as "5," "7," or "9." The names "5," "7," and "9" are used simply as names for their answers. However, "5," "7," and "9" are also used as **names** for the values of the variable we call number and are therefore called numerals. Even though the variable we call "their answer" is not really a number, its values are represented by ________________. — numerals
numerals/numbers

95. Because we said that any variable whose values could be represented by numerals would be called a numerical variable, we would say that the person's answer in the television survey was a

________________ variable, even though the — numerical
numerical/non-numerical

numerals we use as names for the values of this variable ________ represent different numbers in this — do not
do/do not

particular case.

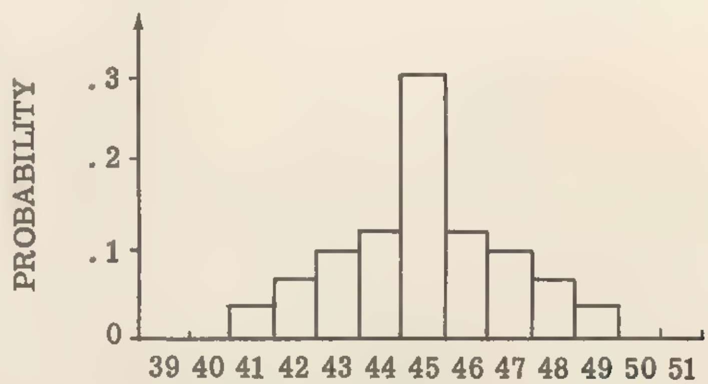
96. Because the list of the **observed** values of the answer variable would be a list of numerals, it would be an example of numerical d______. — data

97. A list of the possible values of answers in the television survey ____________ be called numerical data however. — would not
would/would not

180. (Continued)

attributed to normal and unimportant variations in the materials that make up the parachutes. According to the distribution, the most typical weight for a **non-defective** parachute is ________ ounces. 45

Probability distribution describing the distribution of weights to be expected in a large group of non-defective parachutes.



WEIGHT IN OUNCES

(Rounded off to the nearest ounce)

181. Similiarly, you would never expect to find a non-defective parachute that weighed less than ______ ounces or more 41
than ______ ounces. 49

182. Note the probability that a non-defective parachute will weigh 45 ounces is ________. Therefore, you would .3
expect **about** ______ out of every 100 non-defective 30
parachutes to weigh 45 ounces.

183. If you found a parachute whose weight was **unusually** light or **unusually** ______________, you might suspect heavy
that the parachute was defective.

98. Such a list would not be called numerical **data** because, although it is a list of values represented by numerals (which makes it a numerical list), it is not a list of observed values. Only a list of ____________ values is referred to as **data**. observed

99. Earlier, we listed how many pins a person knocked down each time he rolled a bowling ball. Each time he rolled the ball he could knock down anywhere from 0 to 10 pins. Thus, any number between 0 and 10 was a(n) ________________ value of the variable "pins knocked down." a possible
possible/observed

100. Each value of the variable "pins knocked down" was represented by a numeral. Therefore, "pins knocked down" is an example of a ______________________ variable. numerical
numerical/non-numerical

The numerals 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 are the ________________ values of that variable. possible

101. The list 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 ______________ be considered data because it is simply a list of the ________________ values of the data. would not / possible
would/would not
possible/observed

102. The list of **observed** values of the variable used in the earlier illustration was 0, 0, 2, 10, and 5. This list ______________ be considered data. would
would/would not

177. If you reject the hypothesis that a particular parachute is suitable when it actually is suitable, you would make a Type ________ error, since you are erroneously **rejecting** the hypothesis (one key word). One

178. If you fail to reject the hypothesis that a particular parachute is suitable when it is actually defective, you would be making a Type__________ error. Two
One/ Two

179. If you made a Type One error (**reject** the hypothesis that the parachute is suitable when it is actually suitable), the company will have the unnecessary expense of sending the parachute back to the factory. On the other hand, if you make a Type Two error (**fail** to **reject** the hypothesis that a particular parachute is suitable when it is actually defective), you will send a defective parachute out to be worn by an airman, which could result in the loss of his life. Therefore, in this particular situation, you would be much more anxious to avoid a Type ____________ error, even if this resulted in a higher Two
One / Two
risk of making the other kind of error.

180. Imagine that you were only one of several people who ran tests on the parachutes and your particular task was to weigh the parachute in order to determine if it had an unusual weight, which might indicate defective quality. Imagine the following probability distribution describes the distribution of weights you would normally expect to find among a large group of **non-defective** (suitable) parachutes. The fact that all non-defective parachutes do not have the same weight can be

103. This list of observed values ______________ be would
would/would not
considered **numerical** data, since each value of the variable (the number of pins knocked down) is represented by a ____________. numeral

104. One way in which men differ from one another is whether or not they have a beard. Suppose you were interested in how many men wore beards. If you went to a busy intersection and made a list of whether or not each man who passed wore a beard, your list might look like this after 5 people had passed:

PERSON	BEARD
1	yes
2	no
3	yes
4	no
5	yes

105. Since this list is a record of your observations, it ______________ be referred to as data. could
could/could not

106. Here we simply recorded the values of the variable we were studying as "____" or "____" rather than writing out "He wore a beard" or "He didn't wear a beard." yes, no

174. (Continued)

Suppose you used **a statistical decision-making procedure** to test the following hypothesis.

Hypothesis: The parachute is non-defective (suitable for use). If you ______ (accepted/ rejected) this hypothesis, the parachute could be sent out for use by a pilot, whereas if you ______ (accepted/ rejected) the hypothesis, the parachute would be sent back to the factory for repair.

accepted

rejected

175. You should recall that a Type One error occurs when you erroneously ______ (accept/ reject) the hypothesis and that a Type Two error occurs when you erroneously **fail** to **reject** the hypothesis.

reject

176. A convenient way to remember which is a Type One and which a Type Two error, is to recall that wrongly rejecting a hypothesis is a Type One error and the **single** key word is **rejecting**. On the other hand, erroneously failing to reject is a Type Two error (**two** key words: **failing** and **reject**). Thus, erroneously **rejecting** an hypothesis is a Type ______ (One / Two) error since the **single** (one) key word is **reject**. Erroneously **failing** to **reject** an hypothesis involves **two** key words (______ and ______) and is, therefore, a Type ______ (One/ Two) error.

One

failing, reject

Two

107. When we recorded the observations of whether or not a man wore a beard, we could have used the numerals "1" or "0" instead of "yes" or "no." We could have recorded a "1" for each man who had a beard and a "0" for each man who didn't have a beard. Thus, the list of observations we just presented would look like this:

PERSON	BEARD
1	1
2	0
3	1
4	0
5	(?)

Since the fifth person we saw had a beard, we would complete our list by replacing the question mark with a ____.
1/0 — 1

108. Because we represented the values of the variable we are studying with **numerals**, our data is ____________.
numerical/non-numerical — numerical

109. Earlier, we stated that ____________ is the
number/numeral — number
characteristic of a group or collection of things which we determine by **counting** how many things are in the collection.

110. The numerals "1" and "0" are usually used as names for particular numbers. However, when we use these numerals as names for the two values of the variable "bearded or beardless," they ____________ represent
do/do not — do not
the characteristic we call number.

171. (Continued)

Type One errors (assuming the animal was sick when he really wasn't) than you would be with avoiding Type Two errors (assuming the animal was well when he was really sick).

172. If you were **more** concerned with the risk of infecting other animals in the laboratory than of needlessly destroying your own experimental animal, you would probably be willing to assume that the animal was sick if he ate even a slightly less than normal amount of food during the day on which you based your decision. While this would reduce the risk of erroneously accepting the hypothesis that the animal was well, it would ________________ the risk of deciding the animal was infected when he actually wasn't. — increase
increase/decrease

173. It is obvious that statistical decision-making cannot eliminate all possibility of making an error. What it can do is allow you a choice in determining the risk of the different kinds of error. By choosing a particular decision rule, you can **decrease** the risk of one type of error, even though you may ______________ the risk of making the other kind of error. — increase

174. In some situations, it is obvious that one kind of error would be more costly than the other kind of error. For example, suppose you were in charge of checking the quality of parachutes at a parachute factory. Each parachute that passed in front of your test station could either be defective or non-defective (suitable for use).

111. Thus, we have just seen a case where our data was numerical but the numerals didn't have anything to do with the characteristic we call number. The ________ were simply used as **names** for the different values of the variable we were studying: "1" if he had a beard or "0" if he didn't. — numerals

numerals/numbers

112. When we talk about **number**, we can say such things as "three people are a **greater number** of people than ________ people." — two

two/ten

113. We can say the number represented by the numeral "4" is **twice as large** as the number represented by the numeral "________." — 2

114. We can also say that since 7 minus 5 equals 2, and 3 minus 1 equals 2, the difference between 7 and 5 is the same as the difference between 3 and ______. — 1

115. Sometimes when numerals are used to represent variables other than number, we can make similar statements. For example, it would make sense to say that a temperature of 100 degrees above zero was greater (or hotter) than a temperature of ______ degrees above zero. — 50

50/150

116. It would also make sense to say that "4" pounds was twice as heavy as "_____" pounds. — 2

117. It would also make sense to say that the **difference** in height between a person who was $5\frac{1}{2}$ feet tall and one who was 6 feet tall was the same as the **difference** between someone who was $4\frac{1}{2}$ feet tall and someone who was ______ feet tall, since the difference in height was 1/2 foot in both cases. — 5

5/7

168. A Type One error would result in a costly and unnecessary re-running of the experiment, whereas a Type Two error would only result in the loss of a few additional, inexpensive laboratory animals. Therefore, you would want a **decision rule** that would provide very little risk of a Type ________ error, even if this increased your chances of making the other type of error. — One

(One/Two)

169. Since a Type One error is erroneously rejecting the hypothesis and since a Type Two error is erroneously failing to reject it, you could decrease the probability of a Type ________ error by requiring your observation to be highly unusual (rather than a slightly unusual) before rejecting the hypothesis. — One

(One / Two)

170. On the other hand, requiring the event (observation) to be highly unusual reduces the number of observations that would lead you to **reject** the hypothesis. This means you have ________ the risk of accepting the hypothesis when it is actually false. — increased

(increased/decreased)

171. Imagine that the consequences of the two kinds of errors were different from those we just considered. Suppose it would require little effort to test a new animal were the present monkey destroyed. Suppose also that the other animals in the laboratory were very rare and expensive experimental animals. Therefore, if the infection were allowed to spread, it could be extremely costly for the laboratory. Under these circumstances, you would be ________ concerned with avoiding — less

(more/less)

118. We have seen cases, however, where numerals have been used to represent values of a variable so that statements of this sort are not appropriate. For example, when we considered asking people to choose their favorite from among three television channels, we represented the values of the answer variable with the numerals "5," "7," and "9." It ____________ (would/would not) make much sense to say that choosing Channel 7 was **greater** than choosing Channel 5. — would not

119. Also, it ____________ (would/would not) be appropriate to say that the **difference** between answering "5" and answering "7" was the same as the difference between answering "7" and answering "9" because 7 - 5 = 9 - 7. — would not

120. You should NOT assume that just because a variable is numerical it is similar to "number" in some way. It ________ (is/is not) possible for a variable to be numerical and have nothing more in common with "number" than the use of numerals as names for its values. — is

121. A variable is numerical simply because we use ____________ as names for values of the variable. — numerals

122. Later, we will consider in greater detail the use of numerals to represent variables other than number. For the present, we only wish to emphasize that statements like "greater than," "the same difference as," and "twice as much as" ________________ (are always/may not be) appropriate or make sense — even when a variable is represented numerically. — may not be

123. Many variables of interest to a scientist **are** similar in some way to the variable we call "number." For example, the variable called "length" is similar to "number," so it would make sense to say something ten inches long is **longer** than something ____ (5/12) inches long. — 5

165. The two possible kinds of error that can occur when you decide to accept or reject an hypothesis are:

Rejecting when you really shouldn't (which is called a Type __________ error). One
One/Two

Failing to reject when you should reject (which is called a Type __________ error). Two
One/Two

The more **unusual** you require your observation to be before rejecting the hypothesis, the greater will be the probability of your making a Type __________ error. Two
One / Two

166. Only rejecting the hypothesis when a highly unusual event occurs lowers your chances of wrongly rejecting the hypothesis. On the other hand, since this also reduces the number of observations that would lead you to reject the hypothesis, it usually **increases** your chances of making a Type __________ error. Two
One / Two

167. In order to determine the appropriate decision rule to apply in any particular case, you must consider the relative **cost** of the two kinds of errors. For example, suppose the consequences of making a Type One error in the previous illustration (deciding that the animal **was** sick when he **really wasn't**) would result in the loss of this very important experimental animal and a costly re-running of the experiment. On the other hand, suppose that a Type Two error (deciding that the animal was well when he was really sick) would, at the very worst, result in the contamination of a few more relatively inexpensive animals. Under these conditions, you would probably be more concerned about avoiding a Type __________ error than you would the other type One
One / Two
of error.

124. We determine the **number** of a collection of objects by counting them. We determine the **length** of something in a variety of ways, although the most common procedure is to use a ruler. Both counting and the use of a ruler are procedures by which we determine the appropriate **numeral** to represent a value of a particular variable. Procedures of this sort are called **measurement** procedures. Another example of a m________________ procedure would be the use of scales to determine a person's weight. — measurement

125. The numeral 10 represents a ______________ (larger/smaller) number than does the numeral 5, and ten ounces represents a ______________ (larger/smaller) weight than five ounces. Therefore, it is possible to make statements about values of the variable "weight" similiar to the statement you made about the variable "number." For example, you could say the number 20 is twice as large as the number 10, just as a weight of 20 pounds is ______________ as heavy as a weight of 10 pounds. — larger; larger; twice

On the other hand, some variables are **not** at all similar to the variable called "number," even though you could use numerals to represent values of these variables. For example, the fact that the license number on your automobile is larger than the license number on your neighbor's automobile probably ______________ (does/does not) indicate any difference between the two automobiles (although it might indicate you obtained your license number on an earlier date than did your neighbor). — does not

161. The new **decision rule** requires the animal's food consumption on the test day to be even more unusual for a healthy animal than was required by the old rule, in order to reject the hypothesis that the animal was well, the probability of wrongly rejecting the hypothesis is ______________ in Graph A (which illustrates the new
smaller/larger
rule) than it is in Graph B (which illustrates the old rule). smaller

162. By requiring an even more unusual observation before you reject the hypothesis that the animal was well, you have ________________ the probability of making a
increased/lowered
Type One error (wrongly rejecting the hypothesis that the animal was well), since a healthy animal will consume 0 pellets **less** frequently than "**0 or 1**" pellets. lowered

163. On the other hand, it is perfectly clear that you could eliminate all possibility of making a Type One error by using a **decision rule**, according to which you would never ______________ the hypothesis that the animal
accept/reject
was well, since you then would never reject the hypothesis when it was actually true. reject

164. While you would never erroneously reject the hypothesis that the animal was well if you never rejected the hypothesis, you would have a probability of 1 of failing to reject the hypothesis that he was well when he was actually sick. Stated more formally, while you had reduced the probability of a Type One error to 0, you would have increased the probability of making a __________ __________ error to 1 (you would **always** fail to reject the hypothesis when you should reject it). Type Two

126. Whenever a variable is similar to the variable we call "number," it is possible to assign numerals to represent values of that variable in such a way that you maintain the similarity between that variable and the variable "number." Determining how similar a particular variable is to the variable we call "number," and assigning numerals in the appropriate manner is called **measurement**. We will not have time to consider the topic of **measurement** in this program. The major point we wish to emphasize is that you should be careful **not** to suppose a particular variable is similar to "number" just because that particular variable is numerical. Numerals are often used to represent values of a variable simply because numerals are convenient or familiar names. Just because one value of a variable is represented by the numeral 8 and another value of the variable is represented by the numeral 4 ____________ necessarily imply that one value is in
does/does not
any sense larger or greater than the other one.

127. In this section, we have distinguished between "**numeral**" and "**number**" in order to indicate how easy it is to convert a non-numerical variable into a numerical one by simply substituting numerals for the non-numerical names of the values. On the other hand, in most of your reading you will find no distinction is made between "number" and "numeral." Therefore, throughout the remainder of this program we will use the word "number" without attempting to distinguish between "number" and "numeral." Remember, however, that using "numerals" to represent the values of a variable **does not** insure any similarity between that variable and the variable called "number." The procedure of deciding how similar a variable is to number and assigning numerals in an appropriate fashion to represent this similarity is called m______________. measurement

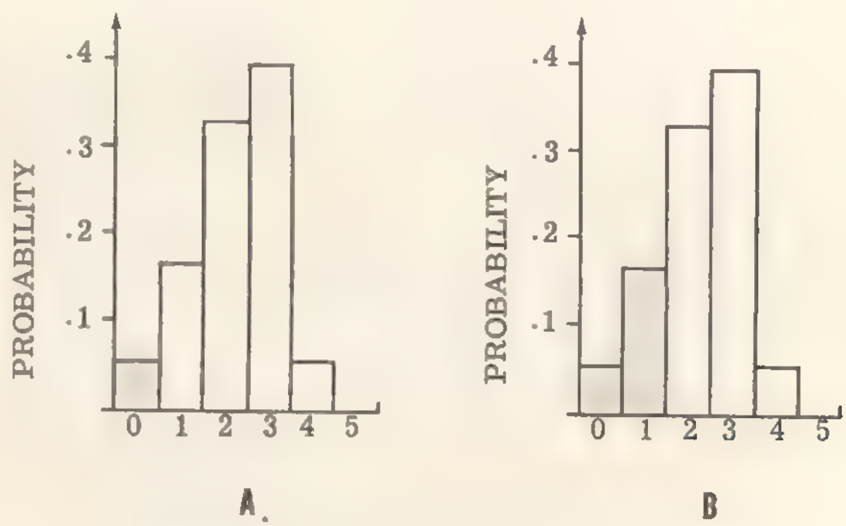
159. (Continued)

represented by Graph ______ (shown below) while the old A
A/ B

decision rule (reject if fewer than 2 pellets are consumed)

could be represented by Graph ______. B
A/ B

Experimental sampling distribution of a healthy monkey's behavior (represented as a probability distribution).



NUMBER OF PELLETS EATEN DURING ONE DAY

160. Once again, the shaded columns in each graph represent the probability of erroneously rejecting the hypothesis that the animal was well, since they represent the probability of a **healthy** animal consuming a number of pellets which would lead you to ____________ the reject
accept/ reject

hypothesis that the animal was well.

128. It is often useful to list things underneath one another, in what is called a **column**. The following list of numerals is arranged in a column.

8

6

5

2

9

The following list of colors is also arranged in a c_______________. column

red

green

blue

green

129. Another way of listing things is side by side, in what is called a **row**. The same list of numerals we just arranged in a column could also be arranged in a row, as follows:

8, 6, 5, 2, 9

The same list of colors we just arranged in a column could be arranged in the following _____________. row
row/column

red, green, blue, green

158. To summarize, then, we have chosen a **decision rule** according to which the probability of making a Type One error is .2. A healthy monkey would eat less than 2 pellets on about 2 out of every 10 days. If you applied the decision rule each day, you would have wrongly decided that the animal was sick on about _______ out of those 10 days. 2

159. One of the most important contributions of statistical **decision-making** is that it is often possible to **specify** the probability of making different kinds of errors. While it is impossible to eliminate the possibility of making errors in many situations, it is very useful to be able to specify the risk — just as it was useful to not only estimate a population parameter but to also specify how accurate the estimate is likely to be.

Furthermore, it is possible to change the decision rule so that the probability of making a particular kind of error is either increased or decreased. This can be illustrated as follows: Suppose you had chosen another decision rule, according to which you would only **reject** the hypothesis that the animal was well if he ate 0 pellets of food on the test day. Thus, this new decision rule could be

130. Since a **column** is a list of things one underneath the other and a **row** is a list of things arranged side by side, these numerals are listed in a __________ and row
row/column

2, 8, 9, 6, 3

these days are listed in a __________. column
row/column

Monday
Tuesday
Wednesday
Thursday
Friday

131. Earlier, we considered an experiment in which we observed how fast a rat ran down an alley to reach food on ten successive days. The following data was presented as an example of what we might have observed:

DAY	RUNNING TIME
1	200 sec.
2	100 sec.
3	150 sec.
4	80 sec.
5	40 sec.
6	41 sec.
7	15 sec.
8	10 sec.
9	4 sec.
10	3 sec.

The different observed values of the running time variable were listed in a __________. column
row/column

154. (Continued)

behavior during the previous 100 days was representative of his **normal** eating habits, the probability distribution shown above represents the eating habits of a monkey who was actually ________.
well/sick

well

155. Since this probability distribution represents the eating behavior of a healthy monkey, the shaded columns represent the probability of ____________ the
accepting/rejecting
hypothesis that the animal is well when he actually is well (when the hypothesis is actually true).

rejecting

156. Since the shaded columns represent the probability of rejecting the assumption that the animal is well when he actually is well, they represent the probability of making a Type One ________.

error

157. You should recall that a Type One error is erroneously rejecting a hypothesis because you had observed an event that was unusual when that hypothesis is actually true. Even the healthy monkey occassionally ate only 0 or 1 pellets in a day, although this was **unusual**. According to the simple addition rule of Probability Theory, you can find the probability of his consuming either 0 or 1 pellet by **adding** the probabilities of each of these outcomes. Our estimate of the probability of eating 0 pellets is .05, and of eating 1 pellet, .15. Thus, the probability of a healthy monkey eating 0 or 1 pellets is ______ + ______, or ______.

.05, .15, .20

132. We can also list things in a pair of columns placed side by side. For example, consider the following:

COLUMN ONE	COLUMN TWO
red	brown
green	orange
blue	pink

Thus, the colors appearing in the first column are red, green, and blue and those appearing in the second column are ___________, _____________ and ______. brown, orange, pink

133. We could also think of the same arrangement of colors as being made up of three rows. For example:

ROW ONE	red	brown
ROW TWO	green	orange
ROW THREE	blue	pink

Thus, the colors appearing in the **first** row are red and brown, while the colors appearing in the _________ row are blue and pink. third

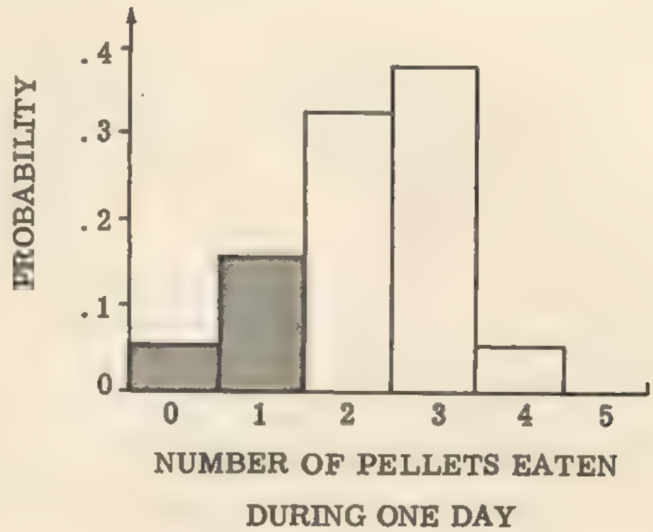
134. A table is an arrangement of things in rows and columns. The arrangement of colors we just considered is a table because it was formed by ____ rows and ____ columns. 3, 2

2/3 2/3

154. (Continued)

This character of the preceding **decision rule** can be illustrated by the following probability distribution.

Experimental sampling distribution of a healthy monkey's eating behavior (represented as a probability distribution).



Notice that we have shaded columns representing the probability of each type of observation that would cause you to __________ (accept/reject) the assumption that the animal is well. However, these probabilities describe the behavior of the **healthy** animal. reject

The unshaded columns represent the probabilities of those observations that would lead us to __________ the hypothesis that the animal was healthy. (Once again, these probabilities describe the eating behavior of a healthy animal.) accept

Notice the preceding probability distribution indicates the probability that the animal will consume any particular number of pellets (as shown along the base of the graph). Since we assumed the monkey's eating

135. The **table** of numerals shown below consists of ______ (2/3) 3

columns and ______ (2/3) rows. 2

	COLUMN ONE	COLUMN TWO	COLUMN THREE
ROW ONE	8	4	9
ROW TWO	6	5	4

Notice in the table of numerals that the numeral 8 is located in the first row AND the first column. The numeral 6 is also located in the first column but in the ______ (first/second) row. second

136. Every numeral in the table falls in a particular combination of row and column. For example, the numeral ______ is located at the intersection of the third column and the first row. 9

The numeral 5 is located in the ______ row and the ______ column. second second

Notice that the numeral 4 is located in the first row, second column and also in the ______ row, ______ column. second third

137. In each of the tables we have seen, "row one" was the row on the ______ (top/bottom) and "column one" was the column on the ______ (left/right). This is the way we will always number the rows and columns in a table. top left

151. The error you would be making in the previous frame would be a Type ________ error. (One/Two) — One

152. If you **reject** an hypothesis when it is really true, you make a Type ________ error. (One/Two) If you **fail** to **reject** an hypothesis when it is really false, you make a Type ________ error. (One/Two) — One; Two

153. Notice that the Type One error in the previous frame has the **single** key word **reject** in bold type. The Type Two error in the previous frame has the **two** key words **fail** to **reject** in bold face type. Thus, to **reject** when you shouldn't is a Type One error since the **single** key word is ____________. To **fail** to **reject** when you should reject is a Type Two error since the **two** key words are ____________ and ____________. — reject; fail, reject

154. One of the most important characteristics of a **statistical decision rule** is that you can often determine the probability of making a Type One error. In other words, you can determine the proportion of times you would expect to **erroneously** reject an hypothesis because an unusual event had been observed which you viewed as inconsistent with this hypothesis.

138. It is often useful to present or record data in the form of a **table.** Suppose you were interested in how a person who took a three-day trip spent his money for food. You could record your observations of this "food cost" variable in a table like the one that follows.

	THURSDAY	FRIDAY	SATURDAY
Breakfast	\$1.00	\$1.15	\$1.50
Lunch	\$2.50	\$3.00	\$1.50
Dinner	\$4.00	\$4.15	\$5.00

Notice that instead of numbering rows and columns, we have identified the ____________ (rows/columns) with the various meals of the day, while the ____________ (rows/columns) are identified with particular days of the trip. The names of the meals and the names of the days identify the names of the data in the columns and the rows. They are not themselves part of the data, since you ____________ (could/could not) have written these row and column headings before you actually made the observations.

rows

columns

could

139. The row and column headings **identify** the data; they ____________ (are/are not) part of the data. Since "breakfast" is the row heading of the first row and "Friday" is the column heading of the second column, the entry in the first row and second column of the data (\$1.15) is the amount spent for ______________ (breakfast/lunch) on ______________ (Thursday/Friday). The amount spent for dinner on Friday was \$______, while the amount spent for __________ on ____________ was \$5.00.

are not

breakfast, Friday

4.15

dinner, Saturday

147. We said that a Type One error was **rejecting** an hypothesis which was really ________ . true
false/true

148. Accepting an hypothesis that was really true or rejecting an hypothesis that was really false ____________ be would not
would/would not
an error.

149. Suppose you tried to use a cigarette lighter and snapped the lever on the lighter several times without lighting the wick. You might decide there was something wrong with the lighter. Even when a lighter is working perfectly, however, you don't necessarily obtain a flame every time you snap the lever. Consider the following hypothesis.

Hypothesis: There is nothing wrong with the lighter.

If the hypothesis is true, then you were simply unlucky the first few times you tried to operate the lighter. If you decided the lighter was broken (or needed lighter fluid) simply because you hadn't ignited the wick, you would be ____________ the previous hypothesis rejecting
accepting/rejecting
when it was actually ________ . true
true/false

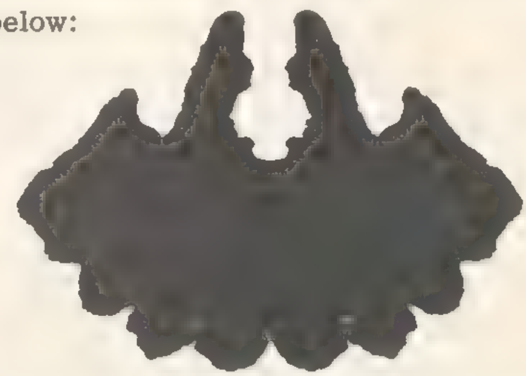
150. If you decided the lighter wasn't working properly when it actually was (simply because you were unlucky and hadn't got a flame the first few times you snapped the lever), you would be ____________ rejecting incorrectly
incorrectly/correctly
the hypothesis that there was nothing wrong with the lighter.

140. How much money was spent for all three meals on Saturday? $__________ 8.00

How much was spent for all three lunches? $_______ 7.00

The day he spent the most money on breakfast, lunch and dinner was _____________. Friday The most expensive meal on each of the three days was ___________. dinner

141. You may have heard of a psychology test called the **Rorschach test**, a test in which people are asked to report their impressions of ink blots, like the one shown below:



This ink blot was formed by spilling ink onto the center of a piece of paper and then folding the paper in half so that the ink is squeezed or blotted between the folds of the paper. When Dr. Rorschach invented this test, he believed the shape of the blot was so vague that a person's answers would reflect as much about the person as about the ink blot. Suppose you were developing a test of this sort. You might make four different ink blots and show all of them to ten people. You could ask each of the people to report whether they thought a particular ink blot gave them an impression of something that was "pleasant," "neutral," or "unpleasant." Thus, their answer would be a

__________________ variable whose three non-numerical
numerical/non-numerical

possible values were "____________________" pleasant

"_______________" and "_______________." neutral, unpleasant

143. If you **fail** to **reject** an hypothesis when the hypothesis is actually **false**, you would be making a Type Two error. Suppose you decided your friend's phone was working perfectly and you had simply been unlucky enough to call him when he was using the phone. If his phone were actually working perfectly, you would be ________ (accepting/rejecting) the hypothesis that it was working perfectly when it actually ________ (was/was not). Therefore, you ________ (would/would not) be making a Type Two error.

accepting

was

would not

144. If you decided that you had just been unlucky and that your friend's phone was really busy when, in fact, it was out-of-order, you would have ________ (accepted/rejected) the hypothesis (the hypothesis that his phone was working perfectly) when you should have ________ (accepted/rejected) it in order to have been correct.

accepted

rejected

145. If you assumed his phone was working properly when it was actually out-of-order, you would have **failed** to **reject** the hypothesis when it was actually false. This would be an example of a Type ________ (One/Two) error. (Remember, a Type One error is **rejecting** when you shouldn't reject, while a Type Two error is **failing** to **reject** when you really should reject.)

Two

146. We said that a Type Two error was **failing** to **reject** an hypothesis when the hypothesis was really ________ (true/false).

false

142. You could record the answers of the subjects in a table which might look like this:

SUBJECT	INK BLOT A	INK BLOT B	INK BLOT C	INK BLOT D
1	pleasant	pleasant	neutral	pleasant
2	neutral	neutral	neutral	pleasant
3	neutral	pleasant	neutral	pleasant
4	neutral	unpleasant	neutral	pleasant
5	pleasant	neutral	pleasant	pleasant
6	unpleasant	unpleasant	unpleasant	unpleasant
7	neutral	neutral	neutral	pleasant
8	pleasant	neutral	neutral	pleasant
9	neutral	neutral	neutral	pleasant
10	pleasant	pleasant	pleasant	pleasant

143. Notice that "INK BLOT A" is a ____________ (row/column) heading and "SUBJECT 1" is a ____________ (row/column) heading.

column

row

144. The row and column **headings** ____________ (are/are not) data, whereas the terms "pleasant," "neutral," "unpleasant" in the table ____________ (are/are not) data (since they are observed values of the "answer" variable).

are not

are

145. The ____________ (row/column) headings identify who made the answer, while the ____________ (row/column) headings identify the ink blot being shown when the answer was made.

row

column

140. (Continued)

On the other hand, if the animal were really sick but happened to eat two or more pellets, you would fail to reject the hypothesis that he was well and you would be making a Type ________ error. Two
One/Two

141. Suppose you tried to telephone a friend several times during the day and you received a busy signal each time. Consider the following hypothesis.

Hypothesis: Your friend's telephone is working properly and the phone was actually busy each time you attempted to call.

If you decided that it was so unusual for his phone to be busy every time you tried to call that you would assume his phone was actually out-of-order, you would be ________________ the previously stated hypothesis. rejecting
accepting/rejecting

142. If you **reject** an hypothesis when the hypothesis is actually **true**, you would be making a Type One error. Suppose you decided your friend's phone was out-of-order when it was actually working perfectly (because you were unlucky enough to call him several times when he was using his phone). You would be ________________ rejecting
accepting/rejecting
the hypothesis (stated in the previous frame) when the hypothesis was __________. Therefore, you true
true/false
______________ be making a Type One error. would
would/would not

146.	All the answers in the first column of the table were made in response to ink blot _____ (A/D), whereas all the answers in row 1 were made by subject _____ (1/10). Subject 1 thought blot A was pleasant, blot B was pleasant, blot C was __________, and blot D was __________. Subject 4 thought blot B was unpleasant, but he thought blot ______ was pleasant. Subject 7 thought blot D was __________.	A 1 neutral pleasant D pleasant
147.	Since the answers from all ten subject to ink blot A are in the first column, we could determine how many subject reported ink blot A was "pleasant" by counting all the occurrences of "__________" in the ________ column.	pleasant, first
148.	Four subjects reported ink blot A seemed "pleasant," whereas __________ subjects reported that ink blot B seemed "pleasant."	three
149.	Subject 1 felt two of the blots were pleasant, whereas subject 2 saw only _______ of the blots as pleasant.	one
150.	What if you were told that one of your subjects was severely depressed and had been under treatment by a psychotherapist. You might expect someone who was very depressed or sad to find _________ (more/less) ink blots "unpleasant" than would a typical subject.	more
151.	You would compare different ___________ (rows/columns) of the table in order to compare how different subjects reacted to the ink blots.	rows

136. Notice that two of these four consequences would represent mistakes (errors), whereas the other two consequences would represent **correct decisions**. For example, outcome 1 would represent a ____________ decision, since we accepted the hypothesis that the monkey was well when it actually was ________.
correct/ incorrect — well/ sick

correct

well

137. The other possible way of being **correct** was to reject the hypothesis that the animal was well when he actually __________ well.
was/ was not

was not

138. The two outcomes of our decision making, which represent **errors** would be to __________ the hypothesis that he was well when he was actually sick, or to __________ the hypothesis that he was well when he was actually well.
accept/ reject — accept/ reject

accept

reject

139. Whenever you attempt to decide whether to accept or reject a particular hypothesis, you run the **risk** of making two kinds of errors. You can __________ the hypothesis when it is really false, or you can ________ the hypothesis when it is really true.

accept

reject

140. Rejecting an hypothesis which is really true because you were unlucky enough to observe a highly unusual event (highly unusual when the hypothesis is true) is often referred to as a **Type One** error. The other sort of error, failing to reject a hypothesis which is really false, is called a **Type Two** error. Thus, if the monkey were actually well but ate such an unusually small number of pellets that you decided he was sick, you would have made a Type ________ error.
One/ Two

One

152. If you were looking for a subject who seemed to find an **unusual** number of the blots unpleasant, you would probably pick subject ______ since this subject found ____________ of the blots unpleasant.

6

all (4)

153. There might be something about particular ink blots which really does make them seem more or less pleasant. If you had to pick out the ink blot which might really appear more pleasant than the rest, you would probably pick ink blot ______,since more subjects found this ink blot pleasant in terms of the other three ink blots.

D

154. Notice in the table of data from the ink blot study how the column headings refer to particular ink blots. We could think of each of these four ink blots as a particular value of a variable, which we might call the "ink-blot" variable. Thus, each column of the table would be identified with a particular __________ of the "ink-blot" variable.

value

155. Similarly, each row is identified with a particular subject. We could think of the subject number as a particular value of a variable we might call the "subject" variable. Thus, a particular row would be identified with a particular _____________ of the "subject" variable.

value

132. Let's suppose that you determine to use the following **rule** for making your decision:

Decision Rule

If the animal eats **fewer than 2 pellets** you will reject the assumption that he is well and destroy him.

In other words, if the number of pellets consumed by the animal today is ______ or ______, you will consider this amount sufficiently unusual for a healthy animal to reject the hypothesis that he is well. 0, 1

133. Let's consider the possible consequences of this type of **decision rule**. First of all, the hypothesis can either be true or not true. In other words, the animal can either be infected or ______ infected. not

134. The two possible decisions we might make on the basis of our observation are to either ____________ or to ____________ the **hypothesis** that the animal is healthy. accept reject

135. Therefore, there are **four possible consequences** of our decision:

1. The animal could be **well** and we could **accept** the hypothesis that he was well.
2. The animal could be **well** and we could ________ the hypothesis that he was well. reject
3. The animal could be **sick** and we could **accept** the hypothesis that he was well.
4. The animal could be ________ and we could reject the hypothesis that he was well. sick

156. When we draw the table, we might include the name of the variable whose particular values are associated with each row. For example, consider the following table:

Subject	INK BLOT			
	A	B	C	D
1	pleasant	pleasant	unpleasant	unpleasant
2	pleasant	neutral	unpleasant	unpleasant
3	unpleasant	neutral	pleasant	neutral
4	neutral	neutral	neutral	unpleasant

In this table, the ink blots can be thought of as a variable, whose particular values are A, B, ______, C

and ______. D

Each value of the ink-blot variable is identified by a particular ____________. column

column/ row

157. You could also think of the subject to whom the ink blot was presented as a variable. Therefore, each particular subject is a particular __________ of that value

variable. The word "subject" in the table we have just seen refers to a variable, and the numerals 1, 2, 3, and 4 identify particular __________ of that variable. values

158. The table is useful in organizing the data in terms of the subject variable and the ink-blot variable. The

____________ are associated with particular ink blots columns

rows/ columns

and the ____________ are associated with the rows

rows/ columns

particular subjects.

128. (Continued)

Notice that the height of each column in the preceding graph represents the ____________ that the animal would consume a particular number of pellets on any randomly chosen day. probability

129. The number of pellets with the largest probability would be a typical number of pellets for the animal to consume on any randomly chosen day. This number of pellets is ______. 3

130. On the other hand, the preceding probability distribution indicates that it would be quite ________ (usual/unusual) for the animal to consume 0 pellets, since the probability of its eating that many pellets is only ______ (.05/.50). In other words, you would expect it to consume 0 pellets of food on only about ______ out of every 100 days. unusual .05 5

131. We expressed the experimental sampling distribution as an estimated probability distribution on a sample space consisting of each possible number of pellets the animal could consume. The problem confronting us is to decide whether the number of pellets he consumes today is so **unusual** (in light of his previous eating habits) that we should ________ (reject/accept) the hypothesis that he is well and destroy him in order to protect the other laboratory animals. reject

159. If you were interested in comparing one subject's responses with another subject's responses, you would compare ____________. On the other hand, if you were interested in comparing the responses to a particular ink blot with the responses to another ink blot, you would compare ____________.
rows/columns

rows

columns

160. For example, in the table we just saw, it is a simple matter to determine **which ink blot** was found unpleasant by the most subjects. Since you are comparing different ink blots, you would compare different ____________. The column that contains the most "unpleasant" responses is _____. On the other hand, to find **which subject** made the most "neutral" responses, you would compare different ____________ (rows/columns). You would find that subject ______ was the one who had made the most "neutral" responses.

columns

D

rows, 4

161. Suppose you were interested in comparing the grades made by four students in three different courses. We could arrange their grades in a table, as follows:

STUDENT	COURSE		
	Math	English	Gym
A. M.	A	A	B
J. T.	B	C	D
A. C.	A	A	D
R. K.	D	B	C

DATA ON GRADES IN VARIOUS COURSES

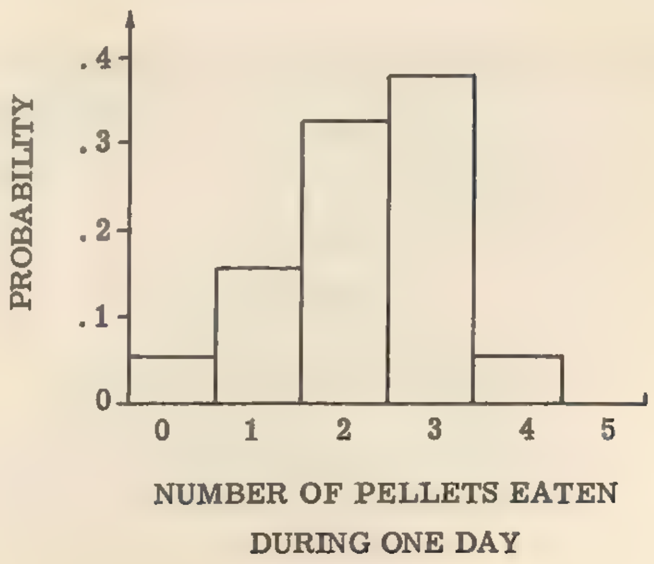
Notice the rows are identified by the initials of particular ____________ (subjects/courses), whereas the columns are identified with particular ____________ (subjects/courses).

subjects

courses

128. Let's assume that the animal's daily food consumption on the previous 100 days **could** be regarded as a simple random process. We could use the proportions in the third column of the previous table as **estimates** of the proportion of days on which the animal would eat each possible number of pellets if we observed him for an unlimited number of days. In other words, we could use each proportion shown in the previous table as an **estimate** of a probability — the probability that the healthy animal would consume a particular number of pellets on any one day. Thus, we could use the following **probability distribution** to represent the animal's eating behavior based upon observations made on the preceding 100 days.

Experimental sampling distribution of a healthy monkey's eating behavior (represented as a probability distribution).



162. The word "course" is written at the top of the table and might be thought of as the name of a variable whose particular values correspond to the particular ____________ (columns/rows) in the table. You could think of ____________, ____________, and ____________ as particular values of the "course" variable.

columns

Math, English, Gym

163. We identify each student by an initial. Thus, if we think of students as a variable, the particular values of that variable are _______, _______, _______, and ______.

A. M., J. T., A. C., R. K.

164. You would find the course in which the students did most poorly by comparing different ____________ (columns/rows). You would find the student who had done the best in his courses by comparing different ____________ (rows/columns).

columns

rows

165. In the preceding table, the variable we are interested in is "grade." The different values of that variable are A, B, C, and D. The data presented in this table are ______ (12/4) observed values of the grade variable.

12

166. Earlier we considered a way in which you might keep track of your improvement as you took bowling lessons. After each lesson, you could bowl 5 balls and record the number of pins you knock down with each ball. For the purposes of this test, you always set up any pins you knock down before you bowl the next ball. Therefore, the fewest pins you could knock down with any one ball are ______, while the greatest number of pins you could knock down with one ball is ______.

0

10

121. The entries in the second column of the previous table indicate on **how many** of the 100 previous days the animal had eaten each of the possible number of pellets shown in the first column. For example, the **first row** of the table indicates that on 5 of the previous 100 days, the animal had eaten _____ pellets. 0

122. On 35 of the previous 100 days, the animal had eaten ______ pellets. 2

123. The **modal** number of pellets consumed daily was _____, 3
since this many pellets had been consumed on ____ days. 40

124. Three pellets is the **modal** number of pellets consumed daily, since this was the **most frequently occurring** daily pellet consumption within the last _______ days. 100

125. In the third column of the above table, we have indicated the _________________ of the previous 100 days on which the animal consumed each of the number of pellets shown in the first column of the table. proportion

126. Since the animal had eaten 0 pellets on 5 of the previous 100 days, we can say the **proportion** of days on which he ate 0 pellets was _______ (as indicated in the third column of the table). .05

127. Since the animal ate 3 pellets on 40 of the previous 100 days, we could represent the **proportion** of days on which he consumed 3 pellets as ________, as the fraction .40
_____/100, or as the percentage _______%. 40/100, 40

167. Suppose you took 4 lessons and bowled 5 of these test balls after each lesson. Imagine you recorded the data from these tests in the following table.

BALL	LESSON			
	1	2	3	4
1	0	0	5	8
2	0	8	4	2
3	3	0	8	5
4	2	2	10	10
5	1	6	6	10

DATA FROM TESTS OF IMPROVEMENT

Since your data are observed values of the "pins-knocked-down" variable, each value is identified with a particular ball you rolled following a particular lesson. In this table, the 5 test balls you rolled are associated with 5 different ____________ (rows/columns), while the 4 lessons that you took are associated with 4 particular ____________ (rows/columns).

rows

columns

168. The total number of pins knocked down following the first lesson was ______ (6/10).

6

169. The total number of pins knocked down following the second lesson was __________. The most pins were knocked down following lesson ______, since there were _____ pins knocked down with the 5 balls rolled on that day.

10
4
35

119. Your decision-making strategy could be summarized as follows: in order to decide whether or not to accept the hypothesis, you will make an **observation** and then decide whether this observation would be such an unusual event if the hypothesis were true, that you would ____________ (accept/ reject) the hypothesis. reject

120. In order to determine what would be an **unusual** amount of food for the monkey to eat in one day, you referred to your records of his previous eating behavior.

Let's suppose your experimental animal ate large pellets of food, which were made available to him each day in the experimental testing chamber. In each of the previous 100 days, he had eaten somewhere between 0 and 5 pellets. The first column of the following table lists the ______________ of pellets the monkey could have eaten on each of these 100 days. number

Eating Behavior of a Healthy Experimental Monkey over a 100 day Period		
Number of Pellets Eaten	Number of Days	Proportion of Days
0	5	.05
1	15	.15
2	35	.35
3	40	.40
4	55	.05
5	0	.00

170. We have now looked at several tables and can summarize certain things about tables. First of all, the definition of a table is "any arrangement of things in ________ and ____________." When the things you are arranging in rows and columns are data, the row and column headings help you to identify each particular piece of data, but they __________ part of the data.
are/ are not

rows, columns

are not

171. It is often useful to think of the row and column headings as particular values of a ____________.

variable

172. One of the first things you should learn to look at when you see a table of data are the row and column headings. These headings identify where the observed values that make up the __________ were obtained.

data

173. For example, the table shown below contains data from two different subjects: subject ____ and subject ____. These two subjects were observed on **four** different __________.

A, B

days

DAY	SUBJECTS	
	A	B
1	10	5
2	6	8
3	9	2
4	4	1

115. (Continued)

other hand, if you ______________ the hypothesis, rejected
accepted/rejected
you should destroy your experimental animal in order to prevent the spreading of the disease to other animals in the laboratory.

116. The problem of deciding whether the animal was ill was a difficult one, since the only noticeable symptom of the illness during the first two days of infection was a **lowering** of the animals' normal food consumption. In other words, if your experimental animal suddenly began to eat less than he normally did, you would probably feel inclined to ______________ the hypothesis reject
accept/reject
that he was healthy.

117. Since the experimental animal was already in the testing chamber, you did not have to worry about his infecting any of the other animals until he was removed at the end of the day. Therefore, you could wait until the end of the day and **observe** how much food he ate that day before deciding whether or not he should be destroyed. If the animal ate an **unusually small amount** of food that day, you could ______________ the hypothesis and destroy the reject
accept/reject
animal.

118. On the other hand, if the amount of food consumed by the experimental animal that day was **not** unusual (i.e., if he ate a normal amount of food), you might be willing to ______________ the hypothesis and continue with the accept
accept/reject
experiment.

174. The data in the table below is ______________________ different from
the same as/different from

the preceding table, since the data below was obtained from four ____________ on two different ________. subjects, days

SUBJECTS	DAYS	
	A	B
1	10	5
2	6	8
3	9	2
4	4	1

This illustrates how carefully you should examine the row and column headings on a table. Many errors in interpreting data can be traced to a simple failure to pay sufficient attention to the headings on the table.

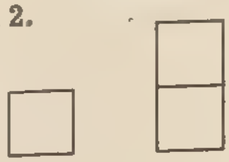
175. You have seen that one way of presenting or recording a list of values for some variable is to arrange the names of these values in ________ and ____________ to form a table. rows, columns

176. Next, we will consider a way of representing values of a variable other than simply arranging the names of the values in rows and columns to form a _________. table

177. If you represented the numeral 1 with a square, you could represent the numeral 2 by putting another square on top of it. Thus, drawing _____ shown below would A
A, B

represent numeral 1, whereas drawing _____ would B
A/B

represent numeral 2.



A B

113. (Continued)

is sufficiently **unusual** to allow you to ________________ reject
accept/reject

the assumption (hypothesis) that the sample came from this particular population.

114. Let's consider another example of the kind of reasoning which takes place in a **statistical hypothesis test**. Suppose you were a psychologist conducting research on monkeys and had used a particular animal in an experiment for over 100 days. One day, while the monkey was in the experimental testing chamber, the animal keeper for the laboratory approached you with a problem. Since several monkeys in the animal colony had contracted a highly contagious disease, it would be necessary to destroy any animals who appeared to have contracted the disease. The problem you faced as an experimenter was to decide whether to ______________ destroy
destroy/save

your experimental monkey (which would disrupt your experiment and waste the data you had already collected) or to ____________ the monkey and run the risk of having save
save/destroy

the disease spread to the other laboratory animals.

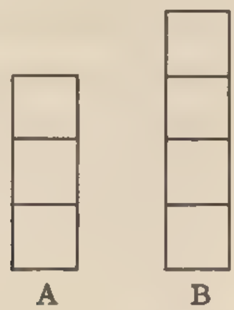
115. In other words, you are faced with the problem of deciding whether to accept or reject the following **hypothesis**:

> The experimental monkey is healthy (he has not contracted the disease).

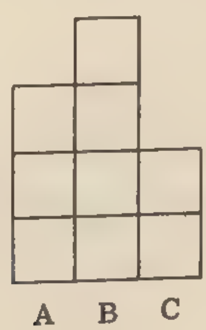
If you ________________ the hypothesis, you could accepted
accepted/rejected

continue using the animal in the experiment. On the

178. By adding a third square on top, you could represent the numeral 3. In fact, you could keep adding squares, one on top of the other, for each new numeral you wished to represent. For example, of the two pictures shown below, Picture A would represent the numeral 3, whereas Picture B would represent the numeral ________. 4



179. You can think of these squares placed atop each other as forming columns, and you could place these columns side by side as shown below. Each column represents a numeral in the same way as the columns of squares you just considered. In the picture shown below, we have identified each column with a letter placed directly underneath the column. Column A represents the numeral "3" since this column is three squares high. Column B represents the numeral "____" since this column is four squares high, and column C represents the numeral "______". 4 2



110. While it would be possible to obtain 100 heads in 100 tosses of a **fair** coin, this sequence of events is so ____________ (likely/unlikely) that you would probably **reject** the idea that the coin was fair. Instead, you would accept the alternative idea that the coin was b______________. — unlikely; biased

111. In each of the previous illustrations, someone was faced with the following type of problem. Would what has been observed be a highly **unusual** event if a particular assumption were true? In other words, if a particular assumption were true, how unusual would it be for what had been ________________ to have occurred? — observed

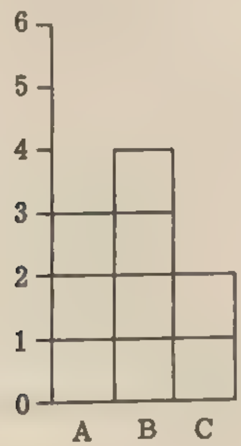
112. If the observation would have been a highly **unusual** event were assumption true, you would tend to ____________ (accept/reject) the assumption. — reject

113. This reasoning procedure can be contrasted with the type of statistical inf____________ we made earlier when we attempted to **estimate** the value of a population statistic on the basis of a sample. Here we make an assumption called an **hypothesis** concerning the population and then ask whether or not the sample we obtained would be so **unusual** a sample from this type of population that we should reject the assumption. You have seen that it is possible to calculate theoretical sampling distributions for random samples obtained from particular types of populations. Since a sampling distribution indicates how often each particular type of sample can be expected to occur, it indicates how **unusual** it would be to obtain a particular type of sample. The sampling distribution indicates whether or not a particular sample — inference

179. (Continued)

The width of each column is the same, but the height differs depending upon which numeral the column represents. In other words, the ____________ of the column determines which numeral it represents. height
width/height

180. The same three columns are shown below. Now, however, we have drawn a line with marks on it to represent the different possible heights of the column. For example, the mark next to the numeral 1 is at the same height as a column one square high. The mark next to the numeral 2 is at the same height as a column two squares high, and the mark next to the numeral 3 is as high as a column _____ squares high. 3



181. Notice Column A is 3 squares high. This is indicated by the fact that it is the same height as the mark next to the numeral _____. 3

107. This problem of choosing between two or more alternative interpretations of a collection of data is a fundamental problem of statistical reasoning. The chief contribution of statistical procedures in reaching decisions of this sort is the ability to specify precisely how unlikely a particular observation would be if a certain interpretation were, in fact, correct. For example, you have seen that it is possible to calculate the probability for obtaining any particular number of heads in a particular number of tosses of a **fair** coin. Suppose you were faced with the problem of deciding whether a particular coin were fair or biased — in other words in deciding whether or not the probability of obtaining heads on each toss of the coin was ______________ or not. Suppose you tossed the coin twice and observed that both times it came up heads. The proportion of heads in your sample would be one, which could lead you to suppose that the coin was biased. On the other hand, the probability of obtaining exactly two heads in two tosses of a coin is one-fourth. Thus, you would expect to obtain two heads in about one out of every ___________ pairs of tosses of a **fair** coin.

one-half

four

108. You would be ______________ (likely/unlikely) to decide that the coin was biased, since two heads in two tosses of a fair coin was a common event.

unlikely

109. However, suppose you tossed the coin 100 times and observed that the coin landed heads each time. This kind of a sample ________________ (would/would not) be **possible** if the coin were fair, even though it would be very unlikely,

would

182. Although Column B is not touching the line with the marks on it, you can see that Column B is 4 squares high and, therefore, at the same height as the mark next to which the numeral _____ is written. 4

183. The mark next to the numeral 2 is at the same height as the top of Column _______ (A/B/C) since that column is two squares high. C

184. Suppose you were tossing a die (one of two dice). You might be interested in the number of dots shown on its top when it came to rest. For example, the following die has _____ dots showing on its top surface. 2

The number of dots appearing on the top of the die after each new toss _____________ (would/would not) be a variable since this number could change or vary from toss to toss. The possible values of this variable could be represented by the numerals "5," "4," "2," "3," "1," and "_____." would; 6

We could represent each possible number of dots showing on the top of the die by a column of squares.

104. (Continued)

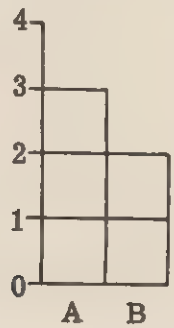
very unusual event. In other words, while it was possible that the television set had failed, you would view this as such a(n) ____________ event, in light of unusual
usual/unusual
your previous experience, that you would **reject** this interpretation in favor of the alternative interpretation that you had been sleeping for a long time and that the television station had gone off the air at its normal time of 2 a.m.

105. Let's consider a third example of the same kind of reasoning process. Suppose you replaced a bad light bulb on your porch. Shortly after you put the new bulb in, it too apparently burned out. Shortly after you replaced the second bulb, a third bulb also burned out. At this point, you decided that it probably was the light fixture itself which was at fault, since it didn't seem reasonable that three light bulbs would all be faulty. In other words, even though there is some possibility that you could have been unlucky enough to have selected three faulty bulbs, you view this possibility as being
____________. You **rejected** this idea in favor of the unlikely
likely/unlikely
alternative assumption that something was wrong with the light fixture itself.

106. In each of the preceding examples, you were faced with the problem of choosing between two alternative interpretations or assumptions. Both of the alternative interpretations were consistent with what you had observed in each illustration. However, you were able to reject one of the two interpretations in each case, on the grounds that what you had ______________ would observed
have been highly unlikely if that particular interpretation had been true.

184. (Continued)

For example, we could represent the die falling so that 3 dots were on top by column ______ shown below: A
A/B



185. Since the largest number of dots we could observe on the top of the die would be _____, the highest possible column of squares we would need would be _____ squares high. 6 6

186. Since the fewest dots we could observe on the top of the die would be a single dot, we could represent this outcome by a column ___ square high. 1

187. Suppose you tossed a die four times and recorded the number of dots showing after each toss, your results might be like those shown in the following table.

TOSS	DOTS SHOWING
First Toss	5
Second Toss	3
Third Toss	6
Fourth Toss	1

103. On the other hand, even if they had fixed the tire, there is some possibility that between the time the tire was fixed and the time you returned that evening, a new injury to the tire had taken place. This second interpretation would be perfectly consistent with what you had seen. However, you probably would reject this interpretation because the number of times that a newly repaired tire would suffer a second flat in such a short while is so ________________ (frequent/infrequent) that it would be more reasonable to assume that the tire had not been repaired. infrequent

104. As another illustration of this kind of reasoning process, imagine that it was the night before your final Statistics examination and that you decided to watch television briefly before studying for the exam. Suppose you suddenly awoke later in the evening, having fallen asleep in front of the television set. Glancing at the screen, you notice that although the set was still turned on, the screen was filled with snow, much as it would if the station had gone off the air. If you had no idea how long you had slept and if there were no clock available, you might immediately draw the conclusion that it was at least 2 a.m., since the television station normally goes off the air at 2 a.m. after the late, late show. However, an alternative interpretation of the same evidence would be that you had not slept that long but rather that something had happened to your set which caused the picture to disappear. **Both** of these interpretations would be **consistent** with the evidence (your observation of the set). However, you would probably **reject** the idea that the television set had failed since this is a

187. (Continued)

On the first toss, there were 5 dots showing, on the second toss there were 3 dots showing, on the third toss there were _____ dots showing, and on the fourth toss there was _____ dot showing. 6 1

188. Notice how the data in the previous table are represented by a ___________ (column/row) of numerals indicating the different observed values of the variable. column

189. The column heading "dots showing" can be thought of as the name of the variable we are observing and the numerals in that column as the _______________ (observed/possible) values of that variable. observed

190. Notice that we have used numerals to represent the different observed values of the variable. These numerals are simply names for the values. Another way of representing these values is with columns of squares that form a sort of **picture** of the data.

Four columns are shown on the following page. Each column represents one of the observed values in the previous table. We have identified the particular toss represented by each column by writing its name directly below that column.

101. When an experimental sampling distribution is expressed in terms of the relative frequency (proportion) of each type of sample, you can think of each proportion as an estimate of the probability of obtaining each particular type of sample. Thus, if you obtained 100 samples by a particular sampling procedure and if you observed that eight-tenths of these samples had the same mean, your best **estimate** of the probability of obtaining that particular type of sample in the future would be ________________. eight-tenths

102. So far, we have considered the problem of **estimating** population statistics on the basis of samples. We have considered which sample statistics provide the most appropriate estimate of the population statistic and also the manner in which you can specify your confidence in the accuracy of the estimate. Next, we shall consider a type of statistical reasoning called **hypothesis testing**, which plays a very important role in research. While some of the more advanced techniques of hypothesis testing can become quite complicated, it is important to realize that the basic procedure of hypothesis testing is almost identical to a very simple reasoning process that you probably use every day.

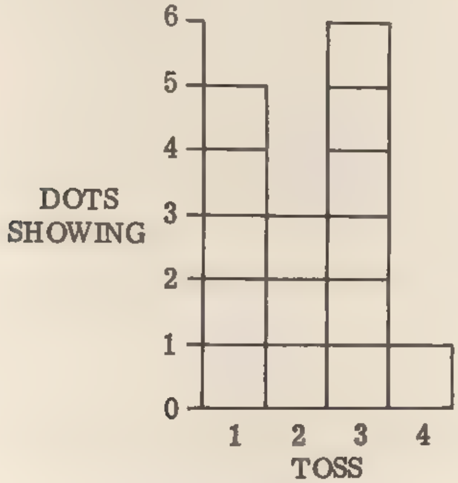
An illustration of this very common reasoning process is provided by the following example. Suppose you left your automobile at a service station in order to have a flat tire repaired. Upon your return later in the day, you notice the car parked in front of the station with the same tire flat. Your obvious conclusion would be that they ______________ fixed the tire. had not

had/had not

190. (Continued)

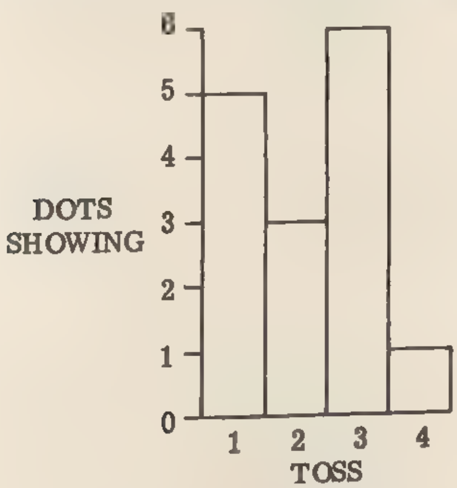
The column for the ________ toss is 6 blocks high since 3rd
3rd/4th

6 dots were showing on the ________ toss. 3rd



191. Since the line next to the column on the left has marks to indicate the different possible heights of the columns, we really don't need the lines between each square in the column.

We have redrawn the columns below, leaving out the lines separating each square in a particular column.



We can still see that the third column represents the numeral 6 since the third column is the same height as the mark next to the numeral ________. 6

97. (Continued)

equivalent to asking what proportion of times would the batter obtain a hit in an unlimited number of times at bat. In this sense, we can view any actual series of times at bat as a ________________ drawn by a random procedure from an infinite population consisting of an unlimited number of times at bat. sample

98. We can view the proportion of hits in any series of times at bat as an **estimate** of the proportion of hits that would have been observed in an unlimited number of times at bat. In other words, we can view the sample proportion as an estimate of a ______________ proportion. population

99. According to the law of large numbers, the more times the player is at bat, the better estimate we will have of the population proportion — since the sample proportion will become more and more similar to the true population proportion. In this sense, a baseball player's batting average (which is simply the proportion of times he has received a hit in previous times at bat) could be viewed as an ______________ of the **probability** of his obtaining a hit each time he came to bat. estimate

100. This is similar to the kind of reasoning behind a statement such as, "The **probability** that it will rain today is .8." This statement really means that it rained on eight-tenths of all the previous days on which conditions were the same as they are today. This means that we view the process which determines whether or not it will rain today as a random process. Our previous observations of this process lead us to estimate the probability of rain as .8 since ______ of all the previous outcomes under these conditions have been "rain." .8

192. A figure of this sort where the values of a variable are represented by the height of each column is a type of **graph.** In this graph, the value of each variable is represented by the ____________ of each column rather than by a name or numeral, as in the previous table. height

193. When you roll a **die**, there will be a certain number of dots showing on its top face when the die comes to rest. For example, the die shown below has ______ dots showing on its top face. 6

194. Suppose you tossed a die four times and recorded the observations shown in the following table:

TOSS	DOTS SHOWING
1st	3
2nd	2
3rd	1
4th	3

According to this data, the **fewest** number of dots were showing on the ____________ toss. third

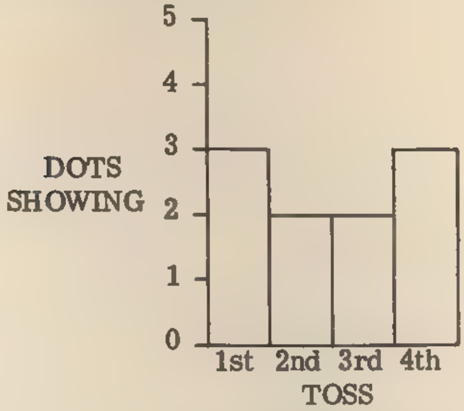
94. Similarly, when we considered drawing a single opinion from the population of student opinions at a particular college, you saw that the probability of drawing a favorable opinion equaled the ________________ of favorable opinions in the population. proportion

95. The proportion of observations having a particular value in a random sample is your best estimate of the proportion of times that value occurs in the population. The probability that a single observation drawn randomly from the population will have a particular value is equal to the proportion of times that value occurs in the population. Therefore, the ________________ of times a particular value occurs in a random sample is your best estimate of the **probability** that a single randomly drawn observation will have that value. proportion

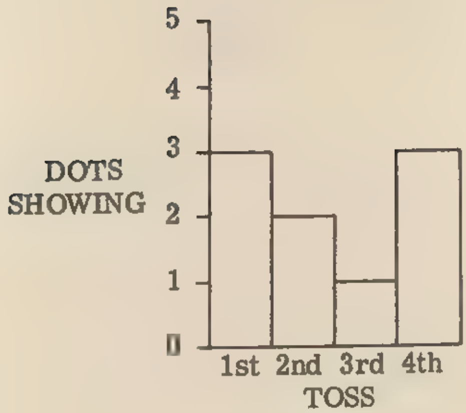
96. In other words, the sample proportion is your best estimate of the **probability** of sampling that particular value from the population. For example, we considered how any series of tosses of a coin could be regarded as a sample from a population consisting of an unlimited number of tosses of a coin. We agreed that if the coin were fair, the proportion of times it would fall heads in an unlimited number of tosses would be ______________. one-half
Therefore, we would say the **probability** of obtaining a head on any single toss was equal to ________________. one-half

97. Similarly, we considered how it might be possible to represent a baseball player's turns at bat as a series of successive outcomes of a random process. If we only considered 2 outcomes (a hit or no hit), we could ask the question, "What is the probability that the batter will obtain a hit on any particular time at bat." This is

195. Two graphs are shown below.



GRAPH A



GRAPH B

In the previous table it was indicated that there were 3 dots showing on the first toss. Both Graph A and Graph B indicate there were ______ dots showing on the first toss because the column representing the first toss in both graphs is ______ squares high.

3

3

196. Compare the data shown in each of the previous graphs with the data shown in the previous table. Graph ______ (A/B) agrees perfectly with the table, whereas graph ______ (A/B) does not.

B

A

90. Even if we supposed that the sampling procedure was random, we still don't know the size of the sample. According to the law of large numbers, you would expect a __________ (large/small) random sample to be more representative of the population than you would a __________ (large/small) random sample.

large

small

91. Variability in sample proportions would tend to be less with large samples than with small samples.

In other words, if the sampling procedure were random, you would expect the variance of the theoretical sampling distribution to be smallest for __________ (large/small) samples.

large

92. If you knew enough about the sampling procedure and the size of the sample to know that the theoretical sampling distribution had a very small variance with most of the sample proportions clustered around the true population proportion, you would have __________ (more/less) confidence in the estimate than you would if you determined that the theoretical sampling distribution had a large variance.

more

93. Earlier, we concluded that if a single observation were drawn randomly from a population, the probability of that observation having a particular value is equal to the proportion of times that value occurs in the population. Thus, if you imagined a population in which half of the observations had the value 8, and if you repeatedly drew a single observation from that population by a random sampling procedure, you would expect about __________ of all the observations sampled to have the value 8.

one-half

196. (Continued)

Both the table and Graph B indicate one dot was showing on the third toss, whereas Graph A indicates ______ dots were showing on the third toss. two

197. Notice how easy it is to compare the values in the two graphs because the values are shown as a **picture**, rather than as numerals in a table. Because it shows a picture of the values rather than simply listing their names, a ___________ is useful. graph
graph/ table

198. Earlier, you considered an experiment in which we recorded the time it took a rat to run down an alley to reach food. The data for that experiment were observed values of the "running-time" variable. Let's consider how we could represent values of the running-time variable in the form of a graph. Four values we might have observed in that experiment are shown below in the form of a __________. table
graph/ table

DAY	RUNNING TIME
1	30 Seconds
2	20 Seconds
3	10 Seconds
4	5 Seconds

87. Whenever you encounter an estimate of a population proportion in your reading, it will be important to consider the probable **accuracy** of the estimate. For example, consider the following statement:

> "Two out of three college students prefer programmed instruction to regular textbooks."

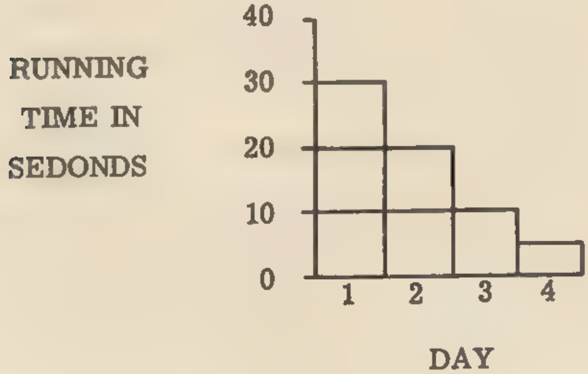
One interpretation of this statement is that two-thirds of all college students in the country prefer programmed instruction to regular texts. It ________ (is/is not) clear, however, whether "2 out of 3 students" refers to the **population** proportion or to a **sample** proportion. — is

88. Another possible interpretation of the statement would be that two-thirds of the students in a sample were in favor of programmed instruction. If this were the case, it would be important to consider the sampling distribution of this sample statistic in order to determine the degree of confidence you should have when you use the sample statistic as an estimate of the population statistic. Furthermore, it ________ (is/is not) clear from the preceding statement whether or not you could use probability theory to calculate a theoretical sampling distribution. — is not

89. One reason it is not clear whether or not you can calculate a theoretical sampling distribution is that you don't know the **procedure** used in obtaining a sample. You could only use Probability Theory to calculate a theoretical sample distribution if the sample procedure were ________________. — random

199. The **longest** observed running time was ______ seconds. 30
The shortest observed running time was _____ seconds. 5

200. The graph shown below represents the same data as the previous table. Notice that we have once again represented a value of the variable by the height of each column. The column representing Day 1 is the same height as the mark next to the numeral 30. The column representing Day 2 is the same height as the mark next to the numeral _______. At the left of the graph we have written what the numerals represent — that is, "running time" in seconds. Since the column for Day 1 is next to a mark with the numeral 30, you know that the rat took ______ seconds to run down the alley to reach the food on the first day. 20 30



201. On Day 4, the rat took 5 seconds to reach the food (as was indicated in the previous table). The column for Day 4 is only half as high as the mark next to the numeral 10. In other words, it represents a running speed of ______ seconds. 5

82. (Continued)

probability of obtaining a sample leading to an error of estimation greater than one-tenth is equal to the sum of the probabilities represented by the ________________ shaded
shaded/unshaded

columns.

83. The probability of making an error of estimation of one-tenth or less is .771. The probability of obtaining any other kind of sample is 1 minus .771, or _______. .229

84. Because of our knowledge of the theoretical sampling distribution, we can make the following confidence statement concerning the accuracy of an estimate. "If the population proportion is one-fifth, the probability of obtaining a sample whose proportion was within one-tenth of the population value is equal to _______." .771

85. You could say the same thing in a different way. You could say:

"Only about ________ out of every 1,000 229
random samples of size 10, drawn from
a population in which p equals one-fifth,
will differ from the true population
proportion by more than one-tenth."

86. The main point of the preceding illustration is that the theoretical sampling distribution indicates the degree of **accuracy** you can expect with a particular size sample randomly drawn from a particular type of population. This allows the person making an estimate to state not only his best bet concerning the population parameter but also his degree of confidence in the a____________ of this estimate. accuracy

202. A table and a graph are shown below. The figure marked A is the ____________ (graph/table), and the figure marked B is the ____________ (graph/table). — table / graph

DAY	SCORE
1	5
2	10
3	7

FIGURE A

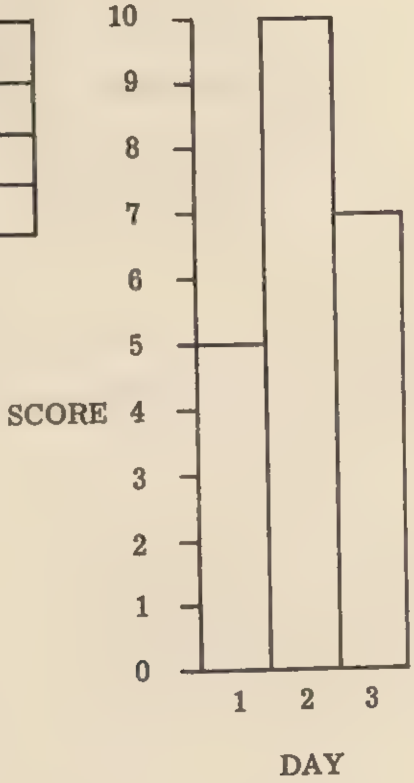


FIGURE B

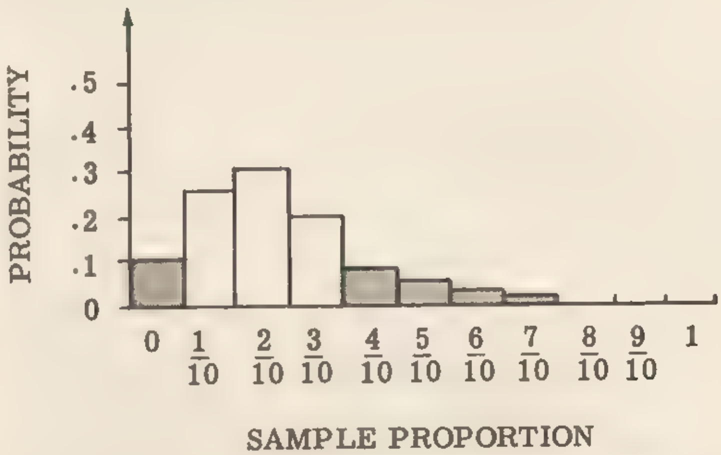
203. The numerals along the side of the graph are different values of the ____________ ("day"/"score") variable. — "score"

204. The largest observed "score" was ______ (10/15) on Day ______ (2/3). Notice how clearly this is shown by the graph: Day ______ (2/3) has the highest column. — 10, 2 / 2

80. There are only 2 possible outcomes in this new sample space. Thus, the probability of obtaining a sample in the other group is simply equal to 1 minus .771, since the sum of the probabilities for the two groups must equal ________. 1

81. We could illustrate the previous theoretical sampling distribution with a graph instead of a table, as shown below.

Theoretical sampling distribution for samples of size 10 drawn randomly from a population in which $\underset{\sim}{p} = \frac{1}{5}$.



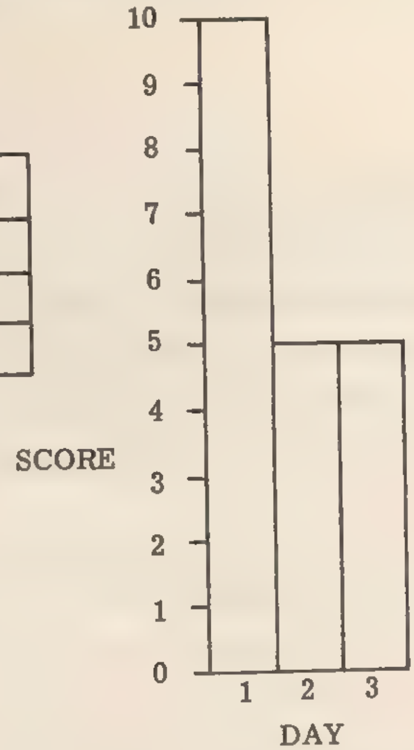
We have shaded the columns representing the probability of obtaining a sample in which the sample proportion would lead you to make an error of estimation greater than ____________. one-tenth

82. The unshaded columns represent the probability of each of the 3 types of samples that would lead to an error of estimation of one-tenth **or less**. Therefore, the probability of obtaining any sample in that group is equal to the sum of these three probabilities. Similarly, the

205. Consider the following graph and table. The "score" value shown for Day 1 ________ the same on the graph is
is/ is not

as on the table.

DAY	SCORE
1	10
2	5
3	?



A score of ____ is shown for Day 2, both on the graph 5
5/8

and on the table.

The observed value for Day 3 is indicated in the

__________ but not in the __________. graph, table
graph/table graph/table

To make the table identical with the graph, you should put a score of _____ in place of the question mark shown 5
in the table.

206. It is clear that the scores were the same on Day 2 and 3 because the columns are the same ____________ for height
both days on the graph.

75. Since the only samples leading to an error of estimation of one-tenth or less are those in which 1, 2, or 3 opinions were favorable, the probability of obtaining a sample in this group is equal to _____ + _____ + _____. .268, .302, .201

76. In other words, the probability of obtaining 1, 2, or 3 favorable opinions in a sample of size 10 is equal to the sum of the probabilities of each type of sample. This sum is equal to .268, + .302 + .201 or _______. .771

77. This means that about 771 out of every 1, 000 random samples of size 10 obtained from this population would be expected to have a sample proportion that differed from the true population proportion by ________________ or less. one-tenth

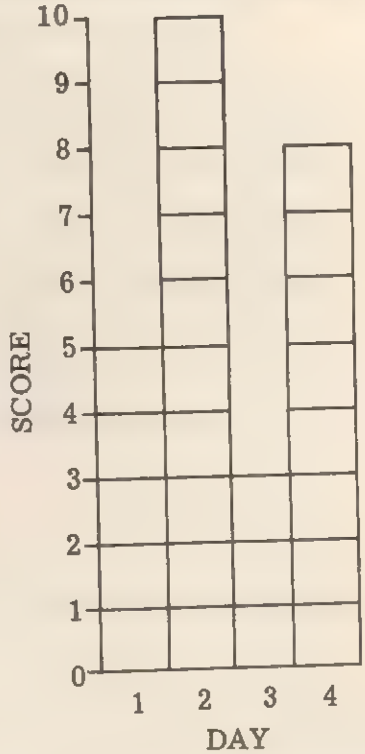
78. The use of s^2 as an **estimate** of the population variance is so common that some authors refer to s^2 as the "sample variance." Naturally, whenever you think of a collection of data as a sample, you are almost always interested in the ___________________ statistics rather than the sample statistics. population

Therefore, you are almost always interested in using the sample to _______________ the population statistics rather than in describing the sample itself. estimate

79. **Strictly speaking,** the variance of the sample (σ^2) is equal to the "sum of the squared deviations from the sample mean" divided by the "number of observations in the sample." You will normally be more interested in s^2, which is equal to the "sum of the squared deviations from the sample mean" divided by the "number of observations minus ________." 1

207. Another graph and table are shown below. Fill in the table so that it is identical to the data shown on the graph. Now compare your work with the table given as the answer.

DAY	SCORE
1	
2	
3	
4	

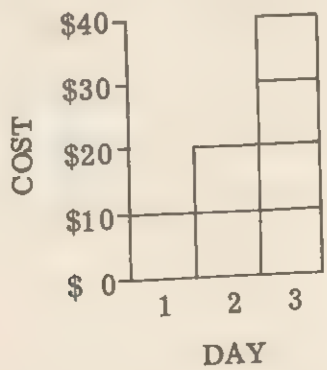


DAY	SCORE
1	5
2	10
3	3
4	8

208. The data in the following graph are values of a variable named "________________." A value of this variable was observed on each of three ______________.

cost

days



70. If the sample contained **less than** _____ favorable opinion or **more than** ______ favorable opinions, the absolute **error** of your estimate would be greater than one-tenth.

1
3

71. The theoretical sampling distribution shown in the preceding table can be regarded as a **probability distribution** on a **sample space** consisting of 11 outcomes. These 11 outcomes are simply the 11 possible types of samples. Each type of sample is assigned a number, or probability, such that the sum of these numbers equals __________.

1

72. In other words, this probability distribution assigns a probability to each of the possible outcomes of the random sampling process. The particular shape of this distribution is a product of not only the random **sampling procedure** but characterizes the **population** as well. Notice that the most frequently occurring types of samples (i.e., the samples with the highest probability) are those in which the proportion of favorable opinions ________ (is/is not) the same as the population proportion ($\tilde{p} = \frac{1}{5}$).

is

73. Notice that samples containing 8, 9, or 10 favorable opinions were so infrequent that when the probability value was rounded off to the closest thousandth it equals _______.

.000

74. Suppose we divided the possible outcomes of the sampling process into two groups: (1) samples that would lead to an estimate of the population proportion which was an error by no more than one-tenth; and (2) all other types of samples. According to the simple **addition rule** of probability theory considered earlier, the probability of obtaining a sample from either group is equal to the ________ of the probabilities of those samples included in the group shown in the previous table.

sum

208. (Continued)

The largest cost was observed on Day _____, since the column for that day is ____________ than any other column.

3

higher

209. To fill in the column and row headings in the following table so that it represented the same things as the graph we just considered, you would label the column containing the values with the word "_________" and the three rows with the numerals "____," "____," and "_____."

cost

1, 2

3

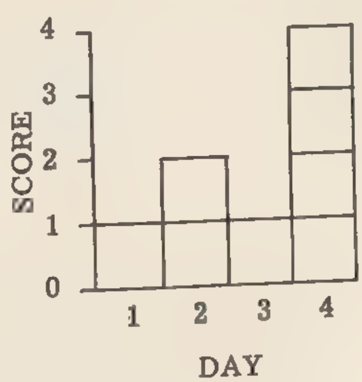
	$10
	$20
	$40

210. Since each row represents a different ________ on which a cost was observed, the square above the numbers 1, 2, and 3 should have the word "_______" written in it.

day

day

211. Fill in the table below so that it represents the same things as the graph does.



DAY	SCORE
1	1
2	2
3	1
4	4

66. Similarly, according to this distribution, you would expect about ______ out of every 1,000 samples obtained in this manner to contain exactly 5 favorable opinions. 26

67. These probabilities represent the **limiting relative frequencies** with which each of the 11 possible types of samples would be expected to occur in an unlimited number of samples obtained in this manner. In other words, as you collected more and more random samples of size 10, you would expect the proportion of samples containing 1 favorable opinion to settle around a value of ____________. By "settle around," we mean that while there would be considerable variability in the proportion of such samples when only a few samples had been collected. As the number of samples collected became very large there would be progressively less variation in the proportion of such samples. .269

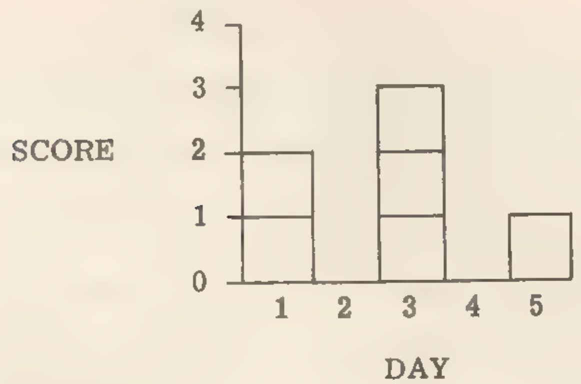
68. Suppose we **grouped** all types of samples into two categories — first, samples that led you to make an estimate of the population proportion which would differ from the true value by no more than one-tenth; second, a group of samples which would lead to an error of estimation **greater than** one-tenth. Since the true population proportion was one-fifth ($\frac{2}{10}$), a sample containing three unfavorable opinions would represent a sample proportion of ______________________. three-tenths

Using this as an estimate would cause you to make an **absolute** error of estimation of ________________. one-tenth

69. An estimate of the population proportion based upon a sample containing one-tenth favorable opinions would be equally as accurate, since the difference between the estimate of one-tenth and the true population proportion (two-tenths) would also represent an absolute error of _______________. one-tenth

212. What if you had a score of zero and you wished to represent it in a graph. If a column three squares high represented a score of 3, and a column two squares high represented a score of 2, and a column 1 square high represented a score of 1, then a column no squares high would represent a score of __________. zero (0)

213. Therefore, on the following graph a score of zero was recorded on Day 2 and Day _______. 4



214. Notice the line on the left of the graph has a mark to indicate the height of each column and a numeral indicating the "score" value of each height. The numeral ____________ indicates that if a column had no height at all it would represent a score of _____. "0" 0

215. Often you will see graphs in which the top of a column is at a height **between** two of the numerals or marks indicating values.

61. (Continued)

This table indicates a theoretical sampling distribution for samples of size 10 drawn randomly from a population in which each member has one of two possible values. For example, it could represent a population consisting of the opinions of all of the students at a university concerning a particular issue, where each student had either a favorable or an unfavorable opinion. We will suppose, furthermore, that the proportion of **favorable** opinions in the population equals one-fifth or, as it is indicated in the table heading, ________ $= \frac{1}{5}$. $\tilde{p}$

62. Every random sample of 10 opinions will lead to one of ________ possible sample proportions. (These are the possible values of p where p equals the proportion of favorable opinions in a sample.) 11

63. Thus, $p = 0$ implies that ________ of the people in the sample had favorable opinions, and $p = 1$ implies that ________ of the people in the sample had favorable opinions. none all

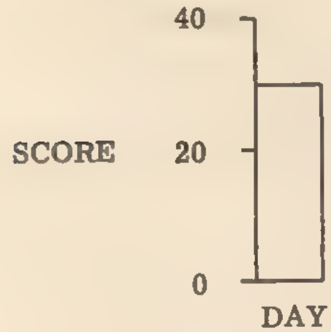
64. We have listed all possible values of the sample proportion (p) in the first column of the table. In the second column we have indicated the probability of obtaining each of the 11 possible types of samples. For example, the probability of obtaining a sample containing 3 favorable opinions is ________. .201

65. Saying the probability of a sample containing 3 favorable opinions is .201 implies that if a very large number of samples were collected, you would expect about 201 out of every 1,000 samples to contain ______ favorable opinions. 3

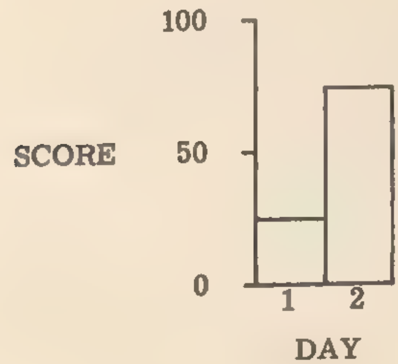
215. (Continued)

For example, the following **column** would probably indicate the value ___________ because it is lower than 30
20/30/40

40 but higher than 20, and approximately half-way between their marks.



216 According to the following graph, the score on Day 1 was ______ and the score on Day 2 was ______. 25, 75
25/50 75/100



59. By the **accuracy** of an estimate, we mean the **absolute** difference between the estimate and the true population statistic. Thus, while a sample mean of 8 would lead to an underestimation of a population mean of 10, and while a sample mean of 12 would lead to an overestimation, both estimates would be equally ______________,since the absolute difference between 10 and 8 is the same as the absolute difference between 10 and 12.

accurate

60. While 12 would represent a positive deviation from the true mean ($\mu = 10$), and while 8 would represent a negative deviation, the absolute size of the deviation would equal _____ in both cases.

2

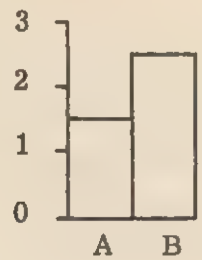
61. If you use a sample statistic as an estimate of the population statistic, the theoretical sampling distribution of the sample statistic indicates how accurate you can expect your estimate to be. For example, consider the theoretical sampling distribution described by the following table.

Theoretical sampling distribution for samples of size 10 drawn randomly from a population where $\tilde{p} = \frac{1}{5}$.	
Sample Proportion	Probability (to nearest $\frac{1}{1000}$)
0	.107
1/10	.268
2/10	.302
3/10	.201
4/10	.088
5/10	.026
6/10	.006
7/10	.001
8/10	.000
9/10	.000
1	.000

217. In the same way, you could represent the value $1\frac{1}{2}$ with a column one and one-half squares high. For example, Column A would represent a score of $1\frac{1}{2}$ and Column B would represent a score of ______ because it is $2\frac{1}{2}$

$2/2\frac{1}{2}/3$

between the height of 2 squares and the height of 3 squares.



218. Suppose you were recording how much time it took a person to solve a puzzle. You might have collected the data shown in the table below:

SUBJECT	TIME TO SOLVE
1	20 Seconds
2	10 Seconds
3	15 Seconds
4	25 Seconds

Thus, there is data from ______ subjects, and each 4

20/4

observed value is a particular ____________. time

time/ subject

55. In some cases, naturally, the descriptive sample statistic is also used as an estimate of the population statistic. Thus, ________ ($\bar{x}$ / μ) is the mean of a sample and it is also used to estimate ________ ($\bar{x}$ / μ). — $\bar{x}$ — μ

56. Similarly, the proportion of times a particular value occurs in a random sample (p) is a(n) ________ (biased/unbiased) estimate of the proportion of times that value occurs in the population ($\tilde{p}$). — unbiased

57. The sample proportion p is an unbiased estimate of $\tilde{p}$, since the **average** value of p in a very large number of samples will tend to equal ________. On the other hand, σ^2 is a biased estimate of $\tilde{\sigma}^2$, since the average value of σ^2 in a large number of samples will tend to be ________ (larger/smaller) than $\tilde{\sigma}^2$. — $\tilde{p}$ — smaller

58. Knowing that a particular estimate is an **unbiased** estimate of a population statistic ________ (does/does not) indicate how accurate an estimate it will be. — does not

By the **accuracy** of an estimate, we simply mean the size of the discrepancy between the estimate and the true population statistic. Therefore, if μ equaled 10, an estimate of 9 would be more ________ than would an estimate of 8. — accurate

1. We have referred to records of the observed values of a variable as **data**.

It is often useful to summarize or describe data, rather than simply listing all the observed __________ which make up the data. values

2. Suppose there were eight possible values of a variable and you observed only three of these possible values. Your data ____________ (could/could not) be summarized by the statement: All the possible values of the variable were observed. could not

3. Of the following two tables, Table ____ (A/B) contains data which might be summarized or characterized by the statement: The value of the variable was the same for all observations. B

TABLE A

Day	Score
1	20 sec.
2	10 sec.
3	20 sec.

TABLE B

Day	Score
1	20 sec.
2	20 sec.
3	20 sec.

Table B indicates the value of the "score" variable was ______ seconds on each of the ______ days. 20, 3

51. However, the fact that s^2 is always ______ (larger/smaller) than σ^2 can be shown to result in a theoretical sampling distribution of s^2 that has a mean equal to $\tilde{\sigma}^2$ (the population variance). — larger

52. The **theoretical sampling distribution** of the sample statistic s^2 has a mean equal to the population variance. In other words, the average value of s^2 computed from a very large number of random samples taken from the same population will tend to equal ______ (μ / $\tilde{\sigma}^2$). — $\tilde{\sigma}^2$

53. The sample variance tends to be an **underestimate** of the population variance, whereas s^2 is just as likely to be an overestimate as it is to be an underestimate. This is why σ^2 is said to be a ______ (biased/random) estimate of $\tilde{\sigma}^2$, whereas s^2 is said to be a(n) ______ (biased/unbiased) estimate of $\tilde{\sigma}^2$. — biased; unbiased

54. It is important to clearly distinguish between sample statistics that simply **describe** the sample and sample statistics that are used as **estimates** of population statistics. For example, ______ (σ^2 / s^2 / $\tilde{\sigma}^2$) is a sample statistic whose primary purpose is to serve as an **estimate** of the variance of the population (i.e., an estimate of ______ (σ^2 / s^2 / $\tilde{\sigma}^2$)). — σ^2; $\tilde{\sigma}^2$

4.

TABLE A

Student	Grade
1	D
2	B
3	C

TABLE B

Student	Grade
1	D
2	A
3	A

You could describe or characterize the data in Table ____ (A / B) above with the statement: A different value was recorded each time the "grade" variable was observed. A

5. Suppose the grades shown in the previous tables were based on the usual grading system of A, B, C, D, and F, where A is the "best" grade and F is the "worst." You could describe the data shown in Table ____ (A/B) with the statement: An A was the best grade observed. B

You could describe the data in Table(s) ________ (A/ B/ A and B) by saying the "worst" grade observed was a D. A and B

6. The reason we say these statements **characterize** or **summarize** the data is that they tell us something, but not everything, about the data. In the previous example, we could describe Table ____ (A/B) by saying: Only two of the possible values of the "grade" variable were observed. This statement describes how many of the possible values are represented in the data. The statement ________ (does/does not) tell us which particular values were observed. B does not

46. Since you are dividing the sum of the squared deviations by a **smaller** number when you calculate s^2 than when you calculate σ^2, it will always be the case that s^2 is larger than σ^2. For example, suppose the sum of the squared deviations from the mean in a sample of size 3 equaled **10**. This would mean that ________ (s^2 / σ^2) equaled 10 divided by **3**, while ________ (s^2/σ^2) equaled 10 divided by **2**, since N equals 3 N - 1 = 2. — σ^2; s^2

47. It is clear that $\sigma^2 = \frac{10}{3}$ is __________ (larger/smaller) than $s^2 = \frac{10}{2}$, since $\frac{10}{3} = \frac{20}{6}$ and $\frac{10}{2} = \frac{30}{6}$. — smaller

48. Notice that the difference between σ^2 and s^2 is trivial when the sample size is very large.

For example, suppose the sum of squared deviations, $\sum(X - \bar{x})^2$ in a sample of size 1, 000 equaled 100. The formula for σ^2 would equal 100 divided by 1000, while the formula for s^2 would equal 100 divided by ________. — 999

49. The **difference** between 100 divided by 1, 000 and 100 divided by 999 is trivial compared to the difference between $\frac{10}{3}$ and $\frac{10}{2}$. This illustrates the fact that the difference between _______ and ________ becomes trivial when the sample size is very __________ (large/small). — σ^2, s^2; large

50. The main reason for using s^2 instead of σ^2 as your estimate of the population variance is that the mean of the theoretical sampling distribution σ^2 is __________ than the true population. In other words, σ^2 typically will be smaller than the true population variance. — smaller

7. Consider the table of data shown below.

Meal	Cost
Breakfast	$1. 50
Lunch	$2. 00
Dinner	$4. 10

The data consist of three values of a variable named

________. cost
"cost"/"meal"

8. You could characterize (describe) the data shown in the preceding table with the statement: The largest observed value of the cost variable was $______. 4. 10

9. Saying $4. 10 was the largest observed value, ____________ does not
does/does not

tell you which particular meal cost the most.

If you were only interested in the smallest amount paid for any one meal, the statement "$1. 50 was the smallest amount paid for any one mean" ______________ describe would
would/would not

the data in sufficient detail for your interests.

10. Therefore, if you were only interested in a particular characteristic of the data, a summarizing statement often answers your question most simply. However, a summarizing statement ____________ tell you as much does not
does/does not

about the data as a complete list or table of data.

42. (Continued)

(Instead of dividing it by the "number of observations in the sample" as you did to compute σ^2.) In other words,

$$\frac{\sum(X - \bar{x})^2}{N} = \underline{\qquad\qquad}$$ σ^2

and

$$\frac{\sum(X - \bar{x})^2}{N - 1} = \underline{\qquad\qquad}$$ s^2

43. Thus, the formula for s^2 differs from the formula for σ^2 simply in that you divide the sum of the squared deviations by N to calculate ________, whereas you σ^2
σ^2 / s^2
divide the sum of the squared deviation by N - 1 to calculate ________. s^2
σ^2 / s^2

44. Remember, to compute σ^2, you divide $\sum(X - \bar{x})^2$ by N, but to compute s^2, you divide by N - 1 instead. Thus, if $\sum(X - \bar{x})^2 = 50$ and $N = 5$, you would find σ^2 by dividing 50 by 5, and you would find s^2 by dividing 50 by ________. 4
5/4

45. Of the following two formulas, formula ________ represents A
A / B
σ^2, while formula ________ represents s^2. B
A / B

A. $\dfrac{\sum(X - \bar{x})^2}{N}$

B. $\dfrac{\sum(X - \bar{x})^2}{N - 1}$

11. One of the first things you might ask about a collection of data is how often were particular values of the variable recorded. For example, the following table indicates a "score" of "8" was observed ______ times out of the 5 observations making up the data. 3

Day	Score
1	8
2	8
3	2
4	5
5	8

The score value 5 occurred the same number of times as the value 2, since they each occurred __________. once

12. You could **characterize** (describe) the data in the previous table by saying the value "_____" appeared three times and the values "_____" and "_____" each appeared once. 8 / 2, 5

This summary of the data ______________ (would/would not) be sufficient if you were only interested in which score value occurred **most frequently**. The statement would tell you that the value "_____" was observed more often than any other value. would / 8

The previous summary statement ____________ (does/does not) tell you enough about the data to determine on what particular day a score of "2" was observed. does not

13. Whether or not a certain way of summarizing the data is suitable ____________ (does/does not) depend on what particular characteristic of the data you are interested in. does

39. We mentioned earlier that statisticians have shown that the variance of a sample is more likely to be **smaller** than the population variance than it is to be larger. In other words, σ^2 is a ______________ (biased/ unbiased) estimate of $\tilde{\sigma}^2$. biased

40. Remember, $\tilde{\sigma}^2$ is the ______________ (population/ sample) variance, population

while σ^2 is the ______________ (population/ sample) variance. sample

41. Statisticians use a sample statistic represented by the symbol s^2 as an estimate of $\tilde{\sigma}^2$. Although s^2 is very similiar to the ______________ (population/ sample) variance, σ^2, it sample

differs just enough to make it a better estimate of $\tilde{\sigma}^2$ than is σ^2.

The chief feature of s^2 making it preferable to σ^2 as an **estimate** of $\tilde{\sigma}^2$ is that s^2 is just as likely to represent as large an overestimate of the population variance as an underestimate. In other words, while the average value of σ^2 based on a large number of samples will tend to be ______________ (equal to/ less than) the population variance, less than

the average value of s^2 based on a large number of random samples will tend to be ______________ (less than/ equal to) equal to

the population variance.

42. The difference in the formulas for σ^2 and s^2 is very slight. In fact, when the sample size is very large, the difference is unimportant. To calculate s^2, you simply divide the sum of the squared deviations from the sample mean by the number of observations in the sample, minus 1.

14. Suppose you asked ten people to judge whether a particular painting was "good" or "bad." You might obtain data of the sort shown in the following table.

Person	Judgement
1	good
2	bad
3	good
4	good
5	bad
6	good
7	bad
8	bad
9	good
10	good

You could summarize this table of data by counting the number of times each of the ____ possible values of the 2
2/3
"judgment" variable were observed.

The two possible values of the variable named "judgment" are ______________ and ______________. good, bad

According to the table of data, the value "good" was observed _____ times and the value "bad" was observed 6
____ times. 4

15. Another way of saying the value "good" occurred 6 times is to say the **frequency** of "good" was 6. Thus, instead of saying the value "bad" occurred 4 times, you would say the frequency of "bad" was _____. 4

36. (Continued)

sampling distributions is that the **average** sample proportion equals the population proportion. In other words, if you obtained many random samples from the same population and calculated the proportion for each sample, the average of all these sample proportions would tend to equal the ______________ ____________. — population proportion

The tendency for the average sample proportion to equal the population proportion is **not** a characteristic of all sample statistics.

It can be shown, for example, that the **variance of a sample** will tend to be slightly **smaller** than the **true population variance**. This does not mean that the variance of every random sample will be smaller than the population variance. It simply means that the sample variance will be ______________ (larger/ smaller) than the population variance more often than it will be ______________ (larger/ smaller). — smaller; larger

37. The variance of a sample is often said to be a **biased** estimate of the population variance, since it tends to be ______________ (larger/ smaller) than the population variance. — smaller

38. You will recall that the variance of a sample is given by the formula shown below:

$$\sigma^2 = \frac{\sum(X - \bar{x})^2}{N}$$

Note that the formula for the variance is simply the sum of the ____________ deviations from the mean divided by the number of observations (N) in the sample. — squared

16. If you were to say the frequency of a certain value was ten, you would simply mean that you counted how often that value had occurred in the data and found it had occurred _____ times. 10

17. If you said the frequency of a certain value was 30, you would mean that you had counted the number of times that value had occurred in the data and found it had occurred _____ times. 30

18. If you had a set of data in which a particular value occurred_____ times, you would say the frequency of that value was 25. 25

19. If 8, 6, 5, 2, 6, and 1 were a collection of data, you would say the frequency of the value 6 was _____ and
2/3 2

the frequency of the value 8 was _____, since "6" occurred twice in the data, whereas "8" occurred only 1

__________. once

20. If a collection of data contained the values 8, 8, 2, 9, 6, 6, and 5, the frequency of the value 8 would be _____ 2
and the frequency of the value six would be _____. 2

21. To say a value has a frequency of zero means that value occurred in the data _____ times. zero (0)

In a particular set of data, therefore, we might find the frequency of the value 20 was **zero**. This would mean the value 20 ____________ in the data. never occurred
never occurred/occurred 20 times

33. Thus, a confidence statement is a useful way of specifying the __________ of making an estimation error of any particular size when estimating a population statistic from a sample. risk

34. In order to specify the risk of obtaining an unrepresentative sample that would lead to a large error of estimation, it is necessary to know how often each type of sample would tend to occur from a particular population with a particular sampling procedure. In other words, it is necessary to know the __________ distribution. sampling

35. Very often you will be interested in estimating the proportion of times a particular value occurs in a population on the basis of a sample from that population. For example, you might wish to estimate the proportion of color-blind people in the United States on the basis of a sample of 100 people drawn randomly from that population. Statisticians show that the proportion of color-blind people in your sample is your best estimate of the population proportion. In other words, if 2 of the 100 people in your sample were color-blind, your best estimate of the proportion of color-blind people in the population would be ________-hundredths. two

36. The main argument for using a sample proportion as your estimate of the population proportion is that this estimation procedure is just as likely to result in an over-estimate of the population proportion as it would an under-estimate. You saw earlier how it was possible to calculate a theoretical sampling distribution of a sample proportion. A characteristic of these theoretical

22. Suppose a variable had three values which could be represented by the three letters A, B, and C. Suppose your data consisted of the following observed values: A, A, B, A, B, B, A, and A. The frequency of the value A is ____ and the frequency of the value B is ___. 5, 3
Since there were no observed values of the possible value C, the frequency of C is _____. 0

23. If the data consisted of ten observations, we would have a list of ten observed values. If all ten of these values were the same, we could say the frequency of that value was _______. ten

24. If the data are 100 observations, it _______________ (would/would not) be possible to have a frequency of some value which was greater than 100. would not

25. Suppose you tossed a coin a hundred times and let it fall on one side or the other each time. The frequency of "heads" in your data could not possibly be greater than _______. 100

26. The fewest number of times "heads" could occur would be none at all. Therefore, the smallest possible frequency of heads would be ______. 0

28. If samples are obtained by a random sampling procedure, statisticians can show that the theoretical sampling distribution of sample means will have a mean equal to the population mean. In other words, they can show that the average or typical value of the sample means will tend to be the same as the ______________ ______________.

population mean

29. Therefore, if you used the mean of a sample as an estimate of the population mean, you will sometimes overestimate or underestimate the mean. However you will tend to over-estimate just about as often as you ______________________.

under-estimate

30. Therefore, one reason for using the mean of a sample to estimate the population mean is that you run the same risk of ov____________________ the population mean as of un____________________ the population mean.

over-estimating
under-estimating

31. It is important to bear in mind that **random** sampling procedures do not guarantee a representative sample. They do allow you to say something about the **risk** of obtaining an unrepresentative sample, since it is usually possible to calculate a theoretical sampling distribution when you use a ________________ sampling procedure.

random

32. For example, while you know that the mean of your sample is your best guess concerning the mean of the population, it would be very useful to know how good a guess it was likely to be. To put it differently, it would be nice to know what risk you ran of making errors of any particular size. A statement that specifies the risk of making an error of estimation of a particular size (or larger) would be similar to the type of c______________ statements we have already considered.

confidence

27. Whenever we count things, we obtain a number. Since we count the values in the data to find their frequency, each frequency ________ (is/ is not) a number. One way of summarizing or characterizing data is to count how often each of the possible values of the variable occur. You could determine a frequency of occurrence for each of the possible values of the variable. — is

Consider the following table of data.

Student	Grade
1	A
2	A
3	B
4	A
5	D
6	F
7	A
8	B
9	B
10	B

The variable represented in the data is named "______" and a value of this variable was recorded for ______ students. Since the grade A occurred 4 times, the frequency of the value A is _____. — grade; 10; 4

28. The frequency of B grades is ______, and the frequency of C grades is ______ because no C's were recorded. D and F both occurred just once. Thus the frequency of D is the same as the frequency of F and equals _____. — 4; 0; 1

23. If samples are obtained by a ____________ sampling procedure, it is often possible to calculate a **theoretical** sampling distribution using the mathematical theory of ____________________.

random

probability

24. Probability theory was originally developed to represent processes such as games of chance (dice, cards, etc.). It is the similarity between the random process involved in a random sampling procedure and the random process involved in games of chance that allows us to use ____________________ ______________ in calculating theoretical sampling distributions.

probability theory

25. Perhaps the chief advantage of using a ______________ sampling procedure is that we can calculate theoretical sampling distributions which indicate how often you would expect to obtain different types of samples from particular kinds of populations.

random

26. Let's consider how useful theoretical sampling distributions can be in drawing inferences about a population on the basis of a sample from that population. One of the most important procedures in ____________________ (inferential/descriptive) statistics is estimating a population statistic on the basis of a sample statistic.

inferential

27. For example, you might wish to estimate __________ (μ / $\bar{x}$) on the basis of _________ (μ / $\bar{x}$).

μ

$\bar{x}$

29. We could summarize this frequency information about the data in the previous table as follows:

Possible Values Of Grades	Frequency
A	4
B	4
C	0
D	1
F	1

The data in this table is represented by the numbers 4, 4, 0, 1, 1. These numbers ________ (do/do not) represent values of the "grade" variable; they do, however, represent the ______________ of times each possible value was observed (occurred in the data). — do not; number (frequency)

30. Because grade A occurred four times, the numeral 4 occurs in the same row of the table as grade _____. — A

31. Because grade B occurred four times, the numeral _______ occurs in the same row as grade B. Since a grade of ________ was never observed, a ______ occurs next to that letter. — 4; C, 0

32. The last two rows in the frequency column contain 1's, since both grade _____ and grade _______ were observed only once. — D, F

33. The table in Frame 27, which contained a grade for each student, is often referred to as a table of the **raw data**. The word "raw" means that we have not summarized the data in any way; we have merely listed all the observations. The data are represented in the

19. One advantage of obtaining samples by a **random sampling procedure** is the fact that as the ______ of the sample is increased, the variability of the sampling distribution tends to become smaller. — size

20. The tendency of a random sample to become more and more representative of a population as the size of the sample is increased is summarized in the law of l______ n______. — large numbers

21. According to the law of large numbers, the frequency of highly unusual or unrepresentative samples will be ______ (larger/smaller) if the sample size is large than it will be if the sample size is small. — smaller

22. Experimental sampling distributions indicate something about the risk involved in making an inference about a population on the basis of a sample. If you only had a single sample from a population, however, you could **not** calculate an experimental sampling distribution. Furthermore, it would often be more difficult to obtain the number of samples necessary to calculate an experimental sampling distribution than it would be to collect all the observations in the population. You have seen, however, that if a sample is obtained in a particular manner, it is often possible to calculate a ______ sampling distribution based on logical considerations rather than on a large number of actual samples. — theoretical

33. (Continued)

table in Frame 29, by the frequency of each value. This table ______ be considered a table of raw data. would not
would/ would not

34. If you added together the frequencies of the possible values shown in the table in Frame 29, you would find that the sum or total of these frequencies equal _____. 10

35. The total of all the frequencies in a frequency table _______ equal to the total number of observations in the corresponding table of raw data. is
is/ is not

36. This is what we would expect, of course, since each value in the table of raw data contributes to only one of the frequencies in the frequency table. For example, the four observed grades of A in the table of raw data were only counted when the frequency for grade _______ was being determined. A

37. If a coin were tossed a hundred times and you were told the observed **frequency** of "heads" was ninety-nine, you would know that the frequency of "tails" was _____ because the frequency of "heads" plus the frequency of "tails" must equal _____. 1 100

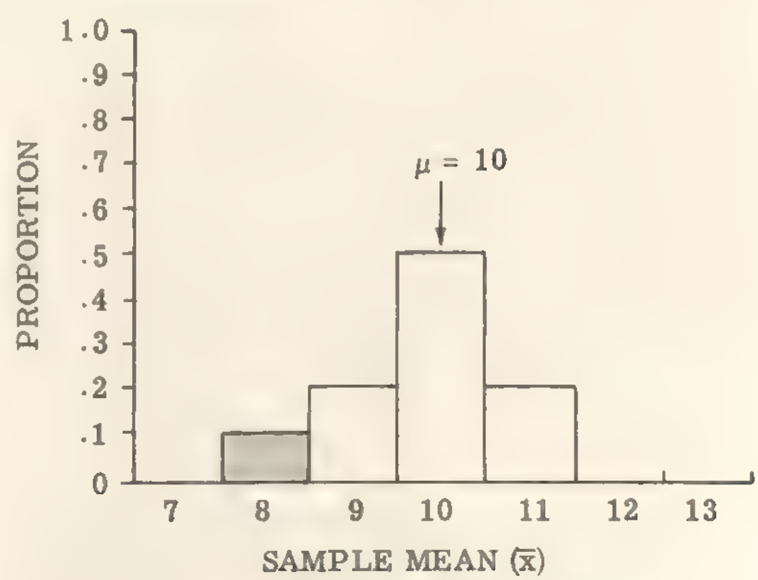
38. If your table of raw data contained 1, 000 observations and a particular value had a frequency of 1, 000, you would know all the other possible values of the variable had frequencies of _____. 0

16. While most of the samples were quite representative of the population, a few samples were less so. Samples with means of 8, for example, would be unrepresentative of the population. Yet, according to the previous distribution, about ________ of all the samples .1
.1 / .5
obtained had means of 8.

17. This experimental sampling distribution indicates something about the **risk** you would be taking when you used a sample mean obtained in this manner to estimate the mean of the population. One way of describing this risk would be by the following **confidence statement**:

"If the mean of the population were actually 10, I would expect to make an error of estimation **greater than** 1 on about only _____/10 of the samples obtained in this manner. 1

18. The fact that about only $\frac{1}{10}$ of the samples obtained in this manner had means differing from the population mean by **more than** 1 is illustrated by the following graph. We have shaded the column representing the proportion of times sample means differed from the true population mean by more than ______. 1



39. Summarizing or characterizing data in the form of a frequency table is suitable if you are only interested in the ______________ of times each value occurred in the data. If you were interested in the sequence or order in which each value was observed, a table of raw data ______________ (would/would not) be suitable for your purposes. — number (frequency) — would

40. Whether a frequency table or a table of raw data is required ______________ (does/does not) depend upon what particular aspect of the data you are interested in. Each of the frequencies in the frequency table is a kind of summary of your data obtained by counting how often each value occurred. We ___________ (can/cannot) think of this frequency as a number that describes or summarizes the data. — does — can

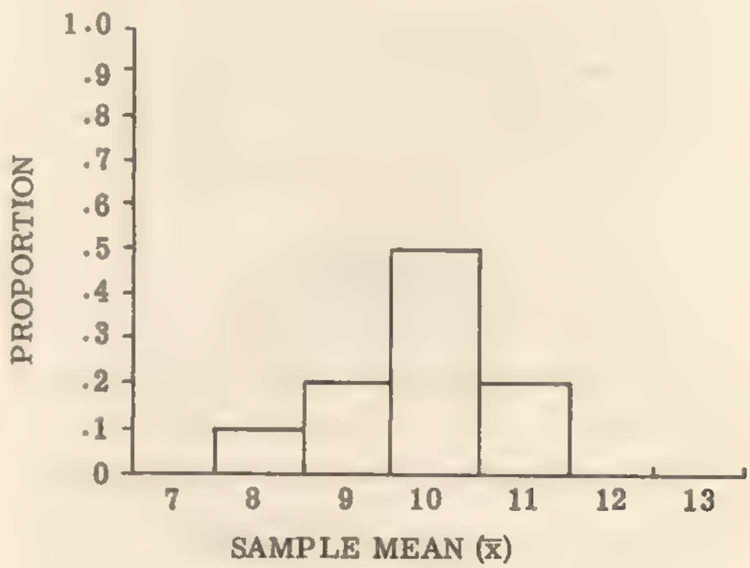
41. Any number or term that summarizes or describes a collection of data is called a **statistic**. Each of the frequencies, therefore, would be called a __________. — statistic

42. Frequencies are often called **enumerative statistics**, because the word enumerate means to count and because we _____________ the number of times a value occurs in the data in order to determine its frequency. Thus, we refer to frequencies as enumerative statistics because the word "enumerate" means to ___________. — count — count

43. Each of the frequencies in a **frequency table** is a number which summarizes the data. Each of these frequencies is called an _________________ statistic because the word "enumerate" means to count. — enumerative

13. An experimental sampling distribution characterizes the sampling procedure, since it suggests how often you should expect to obtain each particular type of sample using a s______________ procedure of that sort to draw samples from that particular population. sampling

14. For example, suppose the proportional frequency distribution shown below were an experimental sampling distribution of sample means obtained by a particular sampling procedure from a population with a mean of 10. (Assume the sample means have been rounded off to the closest whole number.)



15. Notice that the most frequently occurring type of sample had a mean that ____________ (was/ was not) equal to the population mean. was

44. Suppose you were conducting the following experiment. You present a subject with one pair of tones and ask him which of the two tones appeared louder — the first tone or the second tone. Suppose you had presented the subject with ten pairs of tones and asked him make a response after each presentation. You might have recorded the ten responses in the following table of

_______ ___________. raw data

Tone Pair	Answer
1	1
2	2
3	1
4	1
5	2
6	1
7	1
8	2
9	1
10	1

In this table, an answer of "1" represents a response indicating that the first of the two tones was the louder, whereas the answer "2" represents a response indicating that the second of the two tones was the louder. The data in this table could be called, therefore, ____________________ data. numerical

numerical/non-numerical

45. The two possible values of the "answer" variable in the previous table are "____" and "____". The number of observations represented in this table is _____. 1, 2 10

9. There is always some risk involved in drawing conclusions about a population on the basis of a sample, since the sample may or may not be **representative** of the population. The procedure you use for obtaining the sample from a population (along with the character of the population itself) determines how likely you are to obtain a representative sample. You have seen how an experimental ______________ distribution characterizes a particular sampling procedure because it indicates how often each type of sample tends to occur.

sampling

10. Suppose you obtained a large number of samples from a particular population. If you calculated the mean of each of these samples, you could make a frequency distribution of these sample means. This distribution of sample means would be an example of an

______________________ s____________________
experimental/theoretical

distribution of sample means.

experimental
sampling

11. Similarly, you could have calculated some other descriptive statistic, such as the range of each of these samples. The frequency distribution of these sample ranges would be an e________________________

s_________________ d______________________ of

sample _____________.
means/ranges

experimental
sampling
distribution
ranges

12. Suppose you collected 200 random samples of size 5 from a population. You could calculate the median of each sample. The distribution of these _______ medians
200/5

would be an experimental sampling distribution of sample _____________.

200
medians

46. The previous table of raw data could be summarized in the following ______________ table. frequency

Answer	Frequency
1	7
2	3

This frequency table contains the two enumerative statistics "____" and "____." 7, 3

47. Since the total of the frequencies in the frequency table must equal the total number of observations in the table of raw data, we did not have to count the number of times answer 2 occurred if we knew how many times answer 1 had occurred. For example, if the data in the table of raw data had been different and the frequency of the answer 1 had been 6, we would have known immediately that the frequency of the answer 2 was _______, since 10 minus 6 equals 4. 4

48. We have referred to a table containing a list of each observed value as a table of raw data. A table listing the frequency of occurrance of each value is called a ______________ table. frequency

49. Each frequency in a frequency table is the number of times a particular ___________ has been recorded in the data. value

50. Since the frequency of each value is determined by counting, each frequency is a ____________. number

6. **Descriptive statistics** represent one of the two major uses of statistics. The second major role of statistics in research concerns the distinction between samples and populations. As a psychologist, you will often find that while you are interested in the characteristics of a particular collection of data, you only have part of that complete collection available to you. In other words, although you are interested in a ________, you (population/sample) only have a ________ available. (population/sample)

population

sample

7. A statistical procedure for drawing conclusions about a population on the basis of a sample from that population is called an **inferential statistic**. Statistics that simply **describe** a particular collection of data are called ________ statistics. (descriptive/inferential) Another major use of statistics concerns drawing inferences or conclusions about a population on the basis of a sample from that population. This type of statistic is called a(n) ________ statistic. (inferential/descriptive)

descriptive

inferential

8. The most important consideration in drawing conclusions about a population based on a sample is to what degree the sample is **representative** of the population. If you knew all about the population, you could determine to what degree the sample was representative of that population. However, if you knew all about the population, it ________ be necessary to make an (would/would not) **inference** about it based on a sample.

would not

51. A **statistic** is a term or number which describes some characteristic of a collection of data. Thus, each frequency in a frequency table is a ____________. statistic

52. We refer to the frequency of a particular value as an enumerative statistic because the word "enumerate" means to ____________. count

53. Because it lists the value of each observation, the following table could be called a table of __________ data. In other words, the table presents a **complete** list of the data. raw

Observations	Values
1	A
2	B
3	A
4	A
5	A

54. When the observed data contains only two different values, we ___________ (can/cannot) be sure there are only two possible values of the variable. cannot

55. The following table could be used as a ____________ table for the previous table of raw data by filling in the frequencies for each value in the appropriate row of column ____ (1/2). frequency 2

Value	Frequency
A	
B	

1. In the early sections of this program you saw how it is possible to use statistics to **describe** collections of data. For example, a list of the frequencies with which each possible value of a variable occurs in a collection of data is called the d________________ of that variable. — distribution

2. You have seen how the distribution of a variable can be described both in terms of the actual frequencies or in terms of the relative frequencies or p______________. — proportions

3. One of the important features of any collection of values is the typical value or the **central tendency** of the distribution. You have seen how it is possible to represent the central tendency with statistics such as the m__________, m___________, and the m______. — mean, median, mode

4. Another feature of distributions is the degree to which the values are spread out or dispersed. This characteristic of the distribution is often referred to as its variability, which can be represented by statistics such as the v___________ or the r_____________. — variance, range

5. These statistics are often referred to as **descriptive** statistics, since they provide a useful way of _________________ a collection of data. — describing

55. (Continued)

The frequency of the value A in the table in Frame 53 is _______. The frequency of the value B in the same table is _____. Notice how the total of the frequencies in the frequency table ________ (is/ is not) the total number of observations in the table of raw data.

4

1

is

56. Suppose the variable represented in the table of raw data in Frame 53 had four possible values: A, B, C, and D. Since there were ______ observations of the values C and D, the frequencies of C and D would both be _____.

no

0

57. A table of raw data and two frequency tables labeled Table A and Table B are shown below. The frequency table that corresponds to the table of raw data is Table____ (A/ B).

B

Day	Weight
1	5
2	10
3	5
4	10
5	10
6	5
7	5
8	10

Weight	Frequency
5	5
10	3

TABLE A

Weight	Frequency
5	4
10	4
15	0

TABLE B

250. (Continued)

Thus, these probability distributions conveniently represent the relative frequency with which we would expect to obtain various kinds of samples if we repeatedly obtained these samples by a random procedure. We are only able to calculate the form of these theoretical sampling distributions if the samples are obtained by a ______________ process. random

58. Notice that the value $\underset{5/15}{\underline{\qquad}}$ does not appear in the preceding table of raw data and that its frequency ________ presented in Table B. (is/is not) — 15, is

59. Table A could not represent the data in the table of raw data because the value 10 is represented as having a frequency of 3, whereas it should have a frequency of ______, as it does in Table B. Any value of a variable not observed (recorded) in the table of raw data has a frequency of ______. — 4, 0

60. If there were 8 observations in the table of raw data, we would know that the frequency of the value 15 was ____, so long as we knew that the frequency of the value 5 was four and that the frequency of value 10 was four. — 0

61. Every **possible** value of a variable has some frequency — whether or not it is recorded in the data — since a variable has a frequency of ______ if it is not recorded in the data. — 0

62. In summary, we can say a table of raw data lists the ____________ of each observation, whereas a frequency table lists the ________________ of times each value occurred in the table of raw data. — value, frequency (number)

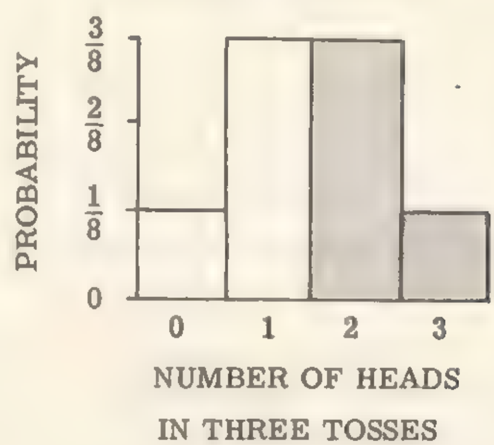
63. Each of the entries in a frequency table is a number and each of these numerals is called an ____________ statistic. — enumerative

The **total** (sum) of these enumerative statistics equals the total number of observations in the table of raw ________. — data

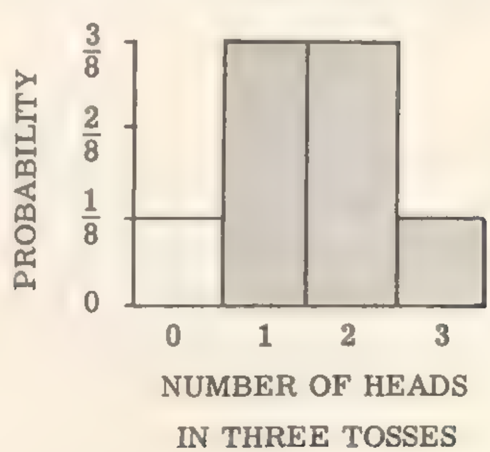
248. Thus, if you wanted to know the probability of obtaining either 2 **or** 3 heads, you would add the probability of obtaining 2 heads and the probability of obtaining 3 heads. In other words:

Pr(2 or 3 heads) = Pr(____) + Pr(_____) 2 heads, 3 heads

249. We have redrawn (below) the previous theoretical sampling distribution, except that we have shaded the columns corresponding to the probability of obtaining _______ **or more** heads in three tosses of a coin. 2



250. On the following graph we have shaded the columns corresponding to the probability of obtaining ____ or more heads. 1



64. A collection of frequencies in a frequency table is called a **frequency distribution** of that variable. A ________________ distribution indicates, therefore, how often each value of a variable occurred in the data. In other words, the collection of enumerative statistics in a frequency table indicating how often each value of a variable occurred is called a f____________ d________________.

frequency

frequency distribution

65. Suppose you asked a subject to sort a collection of ten drawings into four boxes, each box labeled with one of the four adjectives **excellent**, **good**, **fair**, and **poor**. The subject would be distributing the ten drawings amoung the four possible judgments he could make concerning the merit of each drawing. If you recorded his performance, your table of raw data would contain _____ observations of a variable called "judgment,"
10/4
This variable has _____ possible values.
10/4

10

4

66. When the subject was finished distributing the ten drawings among the four boxes, each box would have some particular number of drawings in it. The number of drawings in each box could be thought of as the ____________ of each value of the "judgment" variable. If the subject put five drawings in the box labeled "good" and five drawings in the box labeled "fair," the frequency of "good" judgments would be the same as the frequency of "fair" judgments, and would equal _________.

frequency

5

241. This theoretical sampling distribution indicates the probability of obtaining 0, 1, 2, or 3 heads in _____ tosses of a coin. 3

242. According to this distribution, the probability of tossing a coin three times and getting **no** heads is ________. 1/8

243. Similarly, the probability of obtaining exactly ______ heads in three tosses is also 1/8. 3

244. The probability of obtaining 1 head is the same as the probability of obtaining ______ heads, since it equals __________ in both cases. 2, 3/8

245. One way of interpreting this theoretical sampling distribution is that you would expect to toss a coin three times and get no heads only about ______ out of every 8 groups of three tosses. 1

246. Suppose you were interested in the probability of getting **at least** 2 heads in the three tosses. In other words, you would be asking: "What was the probability of getting either _____ heads or 3 heads in the three tosses?' 2

247. According to the ____________ **rule** of probability theory, the probability of obtaining either 1 outcome **or** another outcome is simply equal to the __________ of their individual probabilities. addition, sum

66. (Continued)

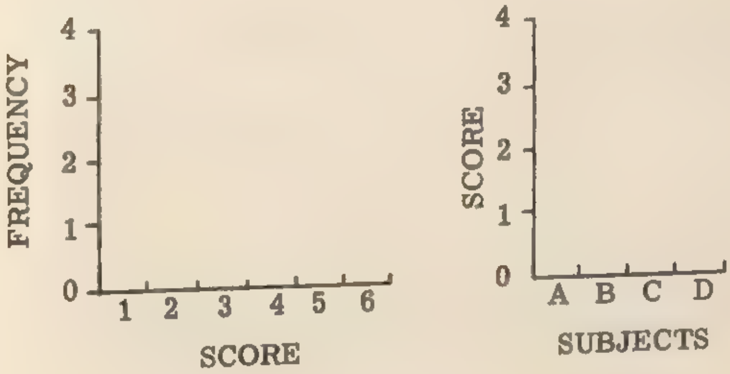
The frequency of "excellent" judgments and of "poor" judgments would both equal ________. 0

The numerals 5, 5, 0, and 0 ____________ (could/could not) be called a collection of enumerative statistics describing the subject's judgments. could

It ____________ (would/would not) be appropriate to say the numerals 5, 5, 0, and 0 in the frequency table define a frequency distribution. would

67. It is often useful to present enumerative statistics by means of a graph. We could represent the frequency of each value with the height of a column, just as we represented the ____________ of a particular observation by the height of a column in earlier graphs. value

Of the two types of graphs shown below, Graph ____ (A/B) would be appropriate for showing the frequency distribution of a variable named "score." A



Graph A Graph B

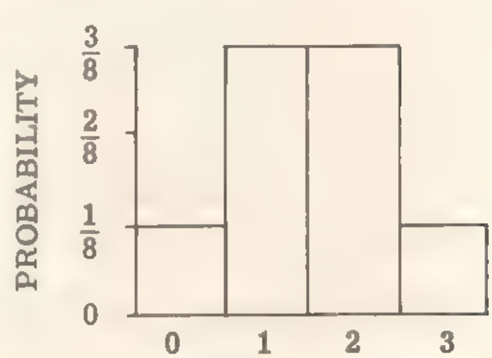
237. Statisticians can show that the mean of a distribution of sample means tends to be the same as the mean of the population. In other words, if μ represents the average height of all of the people in the United States (i.e., the population), then the typical value (average) of the sample means would tend to equal ______. μ

238. Statisticians can also use probability theory to determine a theoretical sampling distribution of sample **means**. It is important, however, to bear in mind that this theoretical sampling distribution is only appropriate for samples drawn in a ______________ or unbiased manner. random

Probability theory only provides a means for calculating theoretical sampling distributions for samples obtained by a ______________ procedure. random

239. The chief advantage of being able to calculate a theoretical sampling distribution is that it tells you how often you should expect to obtain different types of ______________ drawn randomly from a particular kind of population. samples

240. For example, consider the following theoretical sampling distribution. It is an example of a b______________ d________________ of the sort we have already considered. binomial distribution

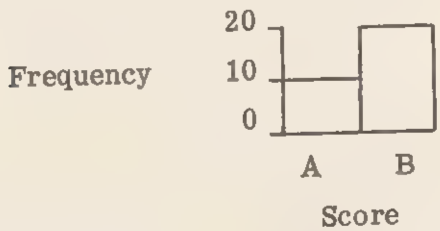


68. If there were only six possible values of the "score" variable, you would know that 6 **columns** would be required in the frequency distribution graph in order to represent all six values. However, since there were only four observations, the largest possible **frequency** for any score would be____. 4

4/6

69. The following frequency distribution graph contains two columns. Score A has a frequency of ______ and 10

Score B has a frequency of _______ (as indicated by the heights of column A and B respectively). 20



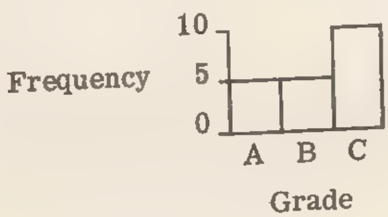
70. The table of raw data for the previous frequency distribution graph would have contained a total of ______ observations. 30

20/30

71. The number of observations represented by the following frequency distribution graph is _____, since 20

A occurred _______ times, B occurred _______ times 5, 5,

and C occurred ______ times. 10



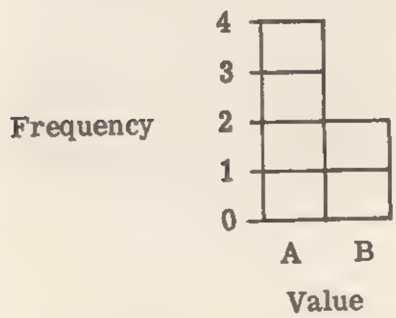
234. The type of continuous curve shown in Figure B (above) is a product of mathematical reasoning, rather than of an actual distribution of data. The important feature of this theoretical distribution is its shape. This can be illustrated quite clearly with a smooth curve, rather than with the series of columns we previously used. Therefore, whenever you encounter a **continuous distribution** represented by a smooth curve rather than a series of columns, you can almost always think of it as if it were made up of the tops of a large number of closely packed ____________. columns

235. We said earlier that **normal distributions** are often used as theoretical sampling distributions, just as binomial distributions are used as ________________ sampling distributions. theoretical

236. Mathematicians have been able to show that if samples are **randomly** drawn from a population with an approximately normal distribution, the distribution of the means of these samples (the sample distribution of sample means) will also have a normal distribution. As an illustration, let's suppose you were to draw many random samples of size 10 from a population consisting of heights of all the people in the United States. Each of these samples of size 10 would have a mean. Therefore, if you drew many samples, you would have many sample means. You can now think of two distributions — first, the distribution of all the heights in the population; second, the distribution of sample ______________. means

72. You know that the most frequently occurring grade in the previous frequency distribution is C, because this grade has the ____________ column. highest

73. You can think of each column in the frequency distribution graph as composed of a series of blocks — one block for each observation of that particular value. For example, consider the graph shown below.



This graph indicates a frequency of _____ for Grade A 4
and a frequency of _____ for Grade B. Therefore, the 2
column for A is four blocks high and the column for B
is _____ blocks high. two

Notice the number of blocks in column A forms a column twice as high as column B, indicating that the frequency of A is ____________ as great as the frequency of B. twice

74. Enumerative or frequency statistics are very important in elections. You are undoubtedly familiar with interpreting the data from elections even if you had not thought of these data as statistics. To illustrate this

232. It is often useful to think of a **continuous distribution,** (represented by the solid curved line as shown in Figure A below), as if it were made up of many tightly packed columns (as shown in Figure B), so that the _________ of the columns formed the shape of the continuous curve shown in Figure A. tops

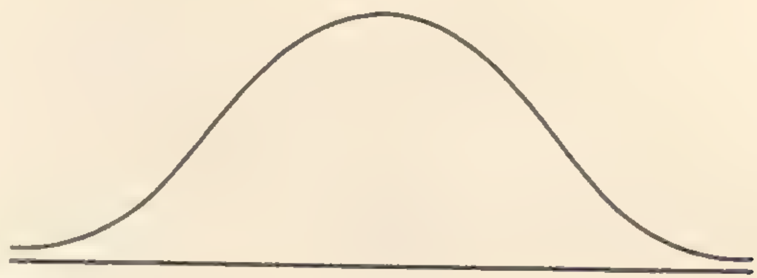


FIGURE A

FIGURE B

233. It is not too important to distinguish between continuous and discrete distributions at this time. We have already shown that a continuous variable is normally treated as if it were _____________. Distributions of a continuous variable are almost always formed by grouping or rounding off the values of the continuous variable. discrete

74. (Continued)

point, suppose that you were conducting an election in which there were three candidates: R. K., A. C. and R. A. We could think of each person's vote as an observation having one of three possible values. These values would be a vote for __________, __________, or __________.

R. K., A. C.
R. A.

75. We could summarize the data from this election in a table such as the one shown below.

Candidate	Number of Votes
R. K.	25
A. C.	75
R. A.	50

Considering "vote" as a variable, this table would be a frequency table indicating that the value "R. K." had a frequency of _____, the value "A. C." had a frequency of _____, and the value "R. A." had a frequency of ____.

25,
75, 50

76. The number of votes cast for candidate R. K. is represented by the number 25 and the number of votes cast for candidate A. C. by the number 75. Both of these numbers are en________________ statistics.

enumerative

77. The group of three numbers, 25, 75, and 50, are the distribution of the 150 votes among the three candidates. In other words, these numbers define a frequency ________________.

distribution

230. On the other hand, using the same data, you could round off the heights to the nearest 5 inches, which would give you ________ (more/less) groups than when you rounded off to the nearest 10 inches. more

231. You have actually divided each of the previous groups into 2 new groups, as shown in Figure 7.

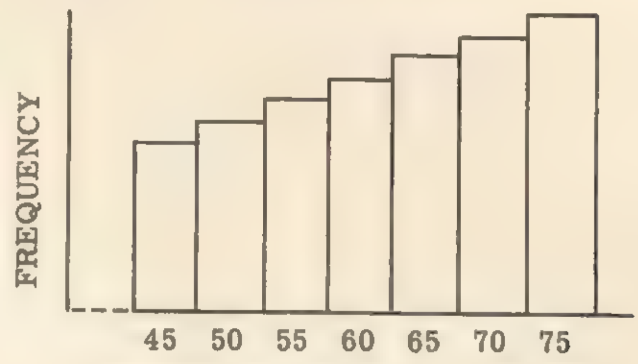
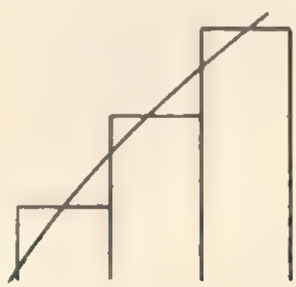
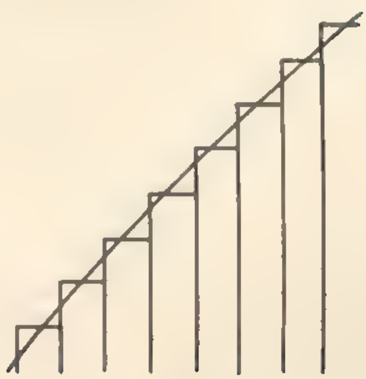


FIGURE 7

Notice that the **shape** of the distribution, as indicated by the **tops** of the columns, is roughly the same, but it has become smoother as you increased the number of columns. For example, notice how the tops of the columns shown in Figure ______ (A/B) (below) are closer to the shape of the smooth curve drawn through them than is so in Figure ______ (A/B). B A

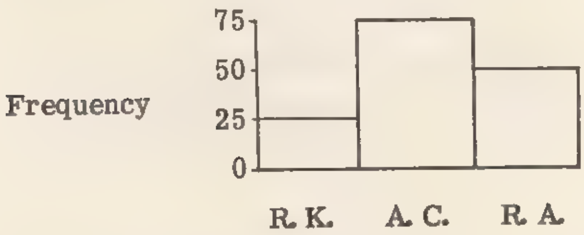


A



B

78. This frequency distribution could also be represented by a graph. The graph shown below _______________ (would/would not) represent the frequency distribution shown in the previous frequency table. — would

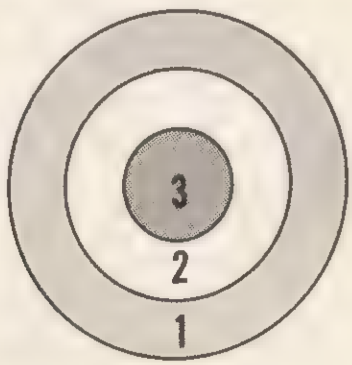


79. Notice the total number of voters in the election could be determined simply by adding all the _______________ in the frequency table. — frequencies (enumerative statistics)

80. It is clear, therefore, that _______ people voted in the election. — 150

81. The winner of the election is indicated by the candidate who has the _______________ column in the frequency distribution graph. This was candidate named ______, who received _______ votes. — highest; A. C.; 75

82. Suppose you were a psychologist interested in studying how accurately a person could throw darts at the dart board shown below:



228. (Continued)

the values "63" or "64", because we are not attempting to distinguish any of the values between these two heights. While many variables can be viewed as **continuous**, we normally treat them as if they were ________________. discrete

229. If you have the heights of 10, 000 students, you could make a frequency distribution in which you rounded off the heights to the nearest 10 inches. For example, in the figure shown on the following page, we have indicated how all of the heights between 46 and 75 inches could be divided into three groups — those between 46 and 55 inches regarded as "50" inches, those between 56 and 65 inches regarded as "60" inches, and those between 66 and 75 inches regarded as "70" inches. In other words, if a person's height were closer to 60 inches than it were to 50 inches, you would call it "___" inches. 60

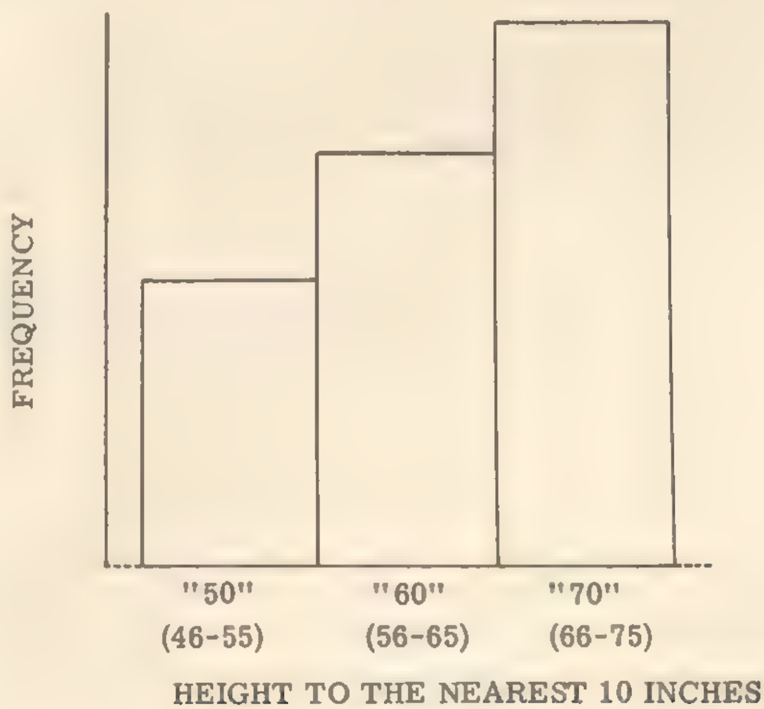


FIGURE 6

82. (Continued)

You might conduct the following experiment. The subject would be asked to stand ten feet away from the target and attempt to hit the innermost ring. Suppose you told him that he would be scored as follows: 3 for the inner ring, 2 for the middle ring, 1 for the outer ring, and 0 for a complete miss. You would then ask him to toss the dart at the board twenty times, recording his score for each toss.

In this experiment, the distance of the subject from the board would be a ______________ (constant/variable), whereas the score the subject received for each toss would be a ______________ (constant/variable).

constant

variable

83. Suppose you obtained the data shown in the following table:

TOSS	SCORE	TOSS	SCORE
1	0	11	1
2	1	12	2
3	1	13	2
4	2	14	2
5	1	15	3
6	0	16	1
7	2	17	3
8	1	18	2
9	3	19	3
10	2	20	3

This table indicates that the subject __________ (did/did not) miss the target completely on his first toss. The first bulls-eye he made was on the ______ toss.

did

9th

226. A normal distribution is an example of a **continuous** distribution, whereas most of the distributions we have considered up to this point are examples of **discrete** distributions. Earlier, in our discussion of the difference between a **continuous** and a **discrete** variable, we indicated that a **discrete** variable was one for which we could locate two values having no other values **between** them. For example, the number of people in a room is a discrete variable since there can be either two people in the room or three people in the room, but there cannot be any number between the values two and three. Thus, "number of people in the room" is a ________ variable. discrete
discrete/continuous

227. On the other hand, a variable such as **height** is a continuous variable. No matter what two pairs of heights we choose, we could always imagine a height between them. For example, one person could be 65.001 inches tall and another could be 65.002 inches tall. It is possible, however, for someone else to have a height of exactly ________, which would 65.0015
65.0015/65.0025
be a value **between** the two previous values.

228. In virtually every case, we treat continuous values **as if** they were discrete by rounding off or grouping values of a continuous variable.

For example, if you rounded off a person's height to the nearest inch, you would simply be **grouping** all the heights that were closer to 63 inches, say, than to 62 or 64 inches and calling those heights "63 inches." Therefore, the variable "height to the nearest inch" only has

83. (Continued)

The subject completely missed the target ________ times out of the twenty tosses. 2

84. To form a frequency distribution, we must count the number of times each ____________ of the score variable occurred in the data. value

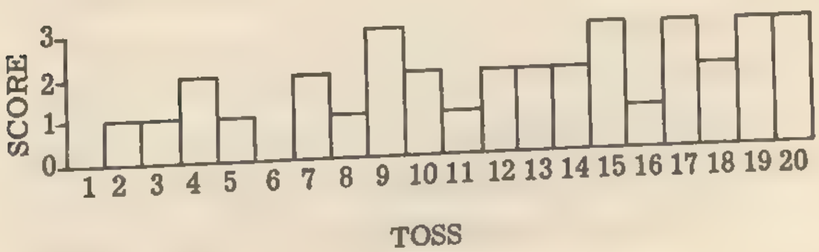
85. We already noted that the value 0 occurred only twice in the data. Therefore, the frequency of a score of 0 would be _______. 2

86. The frequency of a score of 1 is ________. The frequency of 2 is _____ and the frequency of 3 is _____. 6 / 7, 5

87. This group of ____________________ statistics defines a frequency ____________________ of the variable called "score." enumerative / distribution

88. Suppose the data for twenty tosses indicated the frequency of 1's was 5 and the frequency of 2's was 15. The frequencies of both 0's and 3's must therefore be ______. 0

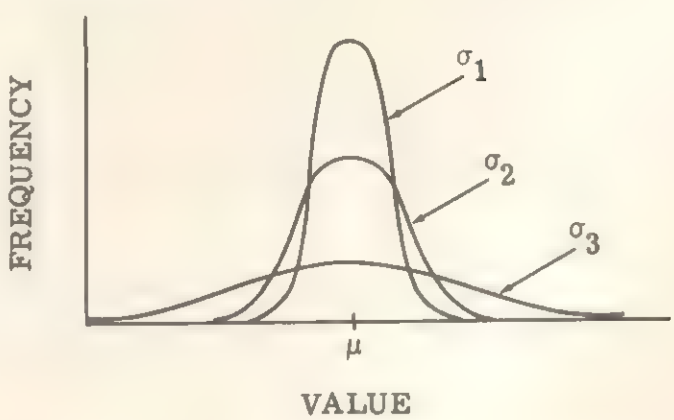
89. A graph of the ______________________________ (raw data/frequency distribution) is shown below: raw data



223. (Continued)

The shape indicated by the dotted line is the characteristic shape of a n__________ distribution. normal

224. This shape, however, may differ considerably, depending upon the standard deviation or variance (since the variance equals σ^2) of the particular normal distribution. The three shapes shown on the following graph are all normal distributions with the same mean (indicated by μ on the base of the graph). These three normal distributions differ only in the size of the variance or standard deviation. Thus, the distribution in which the values are grouped most closely about the mean has a standard deviation indicated by σ $\underset{1\ /\ 2\ /\ 3}{\underline{\qquad\quad}}$. On the other hand, the very flat shape identified by σ_3 is the shape of a normal distribution with a $\underset{\text{large/ small}}{\underline{\qquad\quad}}$ standard deviation. 1 large



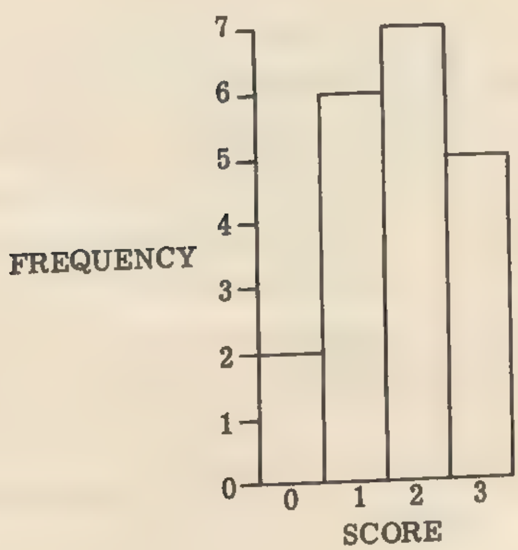
225. Therefore, depending upon whether the standard deviation of the normal distribution is large or small, the actual shape will differ. It is always true, however, that the __________ of values within any particular number of standard deviations from the mean will be the same in any normal distribution, regardless of its particular mean or variance. proportion

90. Graphing the raw data in this way reveals an interesting characteristic of the subject's performance. In general, the more tosses he took, the ______ (better/worse) was his performance. better

91. His performance appears to have improved as he continued his tosses because the columns tend to be ______ (higher/lower) towards the right-hand side of the graph. higher

92. A table and a ______ are shown below. The table contains the frequency distribution we just considered. This same frequency distribution ______ (is/is not) represented by the graph. graph is

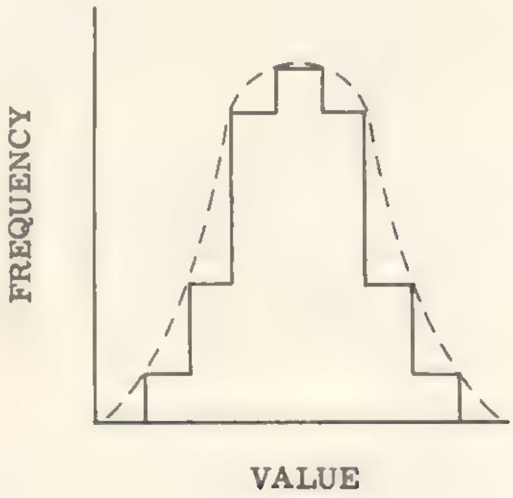
SCORE	FREQUENCY
0	2
1	6
2	7
3	5



222. It is important to note that a normal distribution can have any mean or standard deviation. The important feature of a normal distribution is that the appropriate proportion of values fall within each particular number of standard deviations from the mean. For example, consider two **normal distributions** having a mean of 10.

If one distribution has a standard deviation of 2, then .68 of all the values in that distribution would be between 8 and 12, since any value between 8 and 12 would be less than 1 standard deviation away from the mean. Similarly, if a normal distribution had a mean of 10 but a **standard deviation** of 3, then .68 of all the values in that distribution would fall between the values 7 and ___, 13
since all of these values would be within 1 standard deviation of the mean.

223. Distributions similar to a normal distribution have a characteristic shape often described as **bell-shaped**, such as is shown in the drawing below.

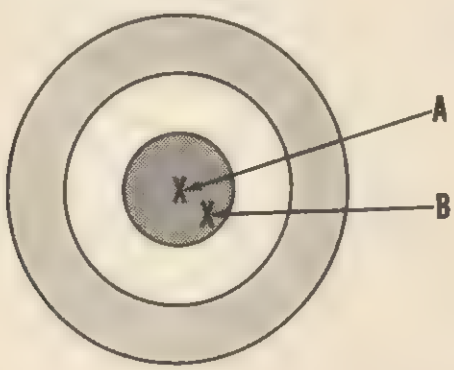


93. The most frequently occurring score is immediately obvious, since the highest column in the frequency distribution graph is for a score of ______. 2

94. While the frequency distribution graph clearly indicates the frequency of each score value, it ____________ (does/does not) indicate the gradual improvement in performance during the course of the experiment. does not

95. A graph of ____________________________ (raw data/frequency distribution) would be the best way to indicate how the subject's performance improved during the course of the experiment, whereas a graph of ____________________________ (raw data/frequency distribution) is the simplest way of showing how the subject's tosses were distributed among the different scores. raw data; frequency distribution

96. In the previous experiment, one toss of the dart might actually be closer to the center of the target than another and yet receive the same score. For example, consider the target shown below:



If the X marks labeled "A" and "B" represent two places where a dart could have hit the target, the point labeled ______ (A/B) would represent a more accurate toss of the dart. A

217. We also noted that standard scores are often referred to as **Z-scores** if the distribution under discussion is a normal distribution. In other words, a value in a normal distribution exactly two standard deviations from its mean would correspond to a Z-score of ________. — 2

218. A Z-score is simply the number of ____________ ______________ between a particular value and the __________ in a ______________ distribution. — standard deviations — mean, normal

219. We have seen how the previous graph can be used to indicate the _________ of values within any number of standard deviations from the mean of a normal distribution. — proportion

220. We say, for example, that .68 of all the values in a normal distribution were within one standard deviation from the mean. This implies that both values less than one standard deviation above and below the mean represent ____________ of all the values in the
.68 / .95
distribution. — .68

221. We can state this same fact in terms of Z-scores by saying that .68 of all the values in a normal distribution correspond to an absolute Z-score of 1 or less. We say **absolute Z-scores** since we are including values with both negative and positive deviations from the mean, so long as the deviation is **equal to or smaller than** ________ standard deviation. — one

96. (Continued)

If all darts falling within the center circle of the target receive a score of 3, the dart hitting at Point A would receive a score of _______ and the dart hitting at Point B would also receive a score of _______. 3 3

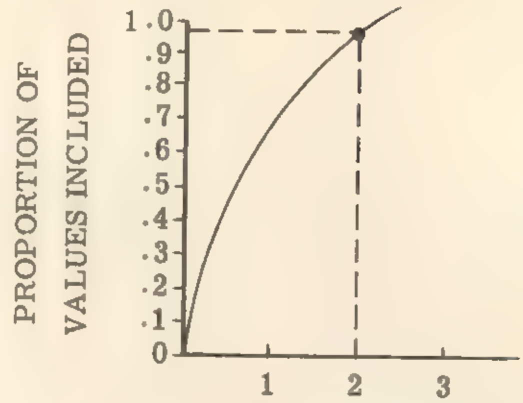
97. Therefore, although one dart was actually thrown more accurately than another, it received the same score. In evaluating the subject's performance, it is easier to **group** all tosses of the dart that hit within the center circle into the same category of accuracy. In a similar manner, we do not consider exactly how far away from the target a dart was if it missed the target completely, since all such darts received a score of _______. 0

98. Another way we might assign a score to each toss of a dart would be to measure exactly how far from the center of the target each dart struck. If we followed this procedure, there _______________ (would/would not) be four possible values of the "score" variable. would not

99. The number of possible values of the "score" variable would depend upon how precisely we wanted to measure the distance between the dart and the actual center of the target.

If we measured this distance to the closest one-thousandth of an inch, there would be ___________ (more/fewer) possible values of the score variable than if we measured down to the closest one-hundredth of an inch. The more precisely we measured the distance, the more possible values of the score variable there would be. more

213. The graph shown below indicates a point on the curve directly above the number ______ on the base of the graph. 2



214. Notice that the point directly above the number 2 on the base of the graph is at the same height as the number ______ on the side of the graph. .98

.98/ .91

215. The fact that the point on the curve directly above the number 2 is at the same height as the number .98 indicates that the proportion of values within two standard deviations of the mean of a normal distribution is ______. You should also recall that a value could be represented as a **standard score**. A standard score was simply the number of standard deviations between a particular value and its mean. We could say that a value two standard deviations away from the mean corresponded to a **standard score** of ______. .98 2

216. Similarly, a value three standard deviations from the mean corresponded to a s______ s______ of ______. If a value corresponded to a standard score of 1, the difference between that value and its mean corresponds to one ______ ______. standard score 3 standard deviation

100. The more precisely we measured the exact distance between the center of the target and the place where the dart struck, the ________ often would we find two tosses of exactly the same score. If we measured closely enough, in fact, we would probably find that the subject never received exactly the same score on any two tosses. less

more/less

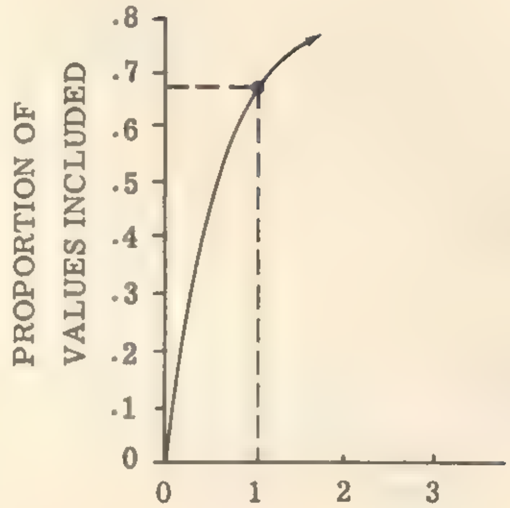
101. It is often useful to consider two observations that are really slightly different as having the same value. In other words, to **group** together observations that are sufficiently similar and to consider them as having the same value is often a useful procedure. The following illustration will indicate how we can group observations in this way.

Suppose you were studying how long it took a person to solve a certain mathematical problem. If you made observations on **twenty** people, recording the time it took each person to solve the problem, your data might be represented by the following table:

Subject	Time in Seconds	Subject	Time in Seconds
1	60	11	20
2	128	12	72
3	68	13	71
4	69	14	120
5	80	15	85
6	75	16	116
7	32	17	99
8	149	18	15
9	16	19	36
10	110	20	86

211. (Continued)

To determine the proportion of values in a normal distribution within **one** standard deviation of the mean, you simply locate the number 1 on the bottom of the graph and then locate a point on the curve directly **above** the number 1, as indicated on the graph shown below.



Notice that the point on the curve directly **above** the number 1 on the base line is at the same height as the number ______ on the side of the graph (as indicated .68
.68/.78
by the horizontal dotted line).

212. The fact that .68 is at the same height as a point on the curve directly above the number 1 on the base of the graph indicates that exactly .68 of all the values in a normal distribution are within one standard deviation of the mean. To find the number of values that are within two standard deviations of the mean in a normal distribution, you first locate a point on the curve directly above the number ______ on the base of the graph. 2

101. (Continued)

Notice there ________ a value of the time variable is not
is/is not

which occurs more than once in the table of data. (You might wish to insert a bookmark indicating this page since we will refer to the table shown above in the next several frames.)

102. Since every observed value of the data occurs once and only once, each of these observed values would have a frequency of ______. 1

103. Notice that the highest value (the longest time) recorded was _______ seconds, whereas the lowest 149
value recorded was _______ seconds. 15

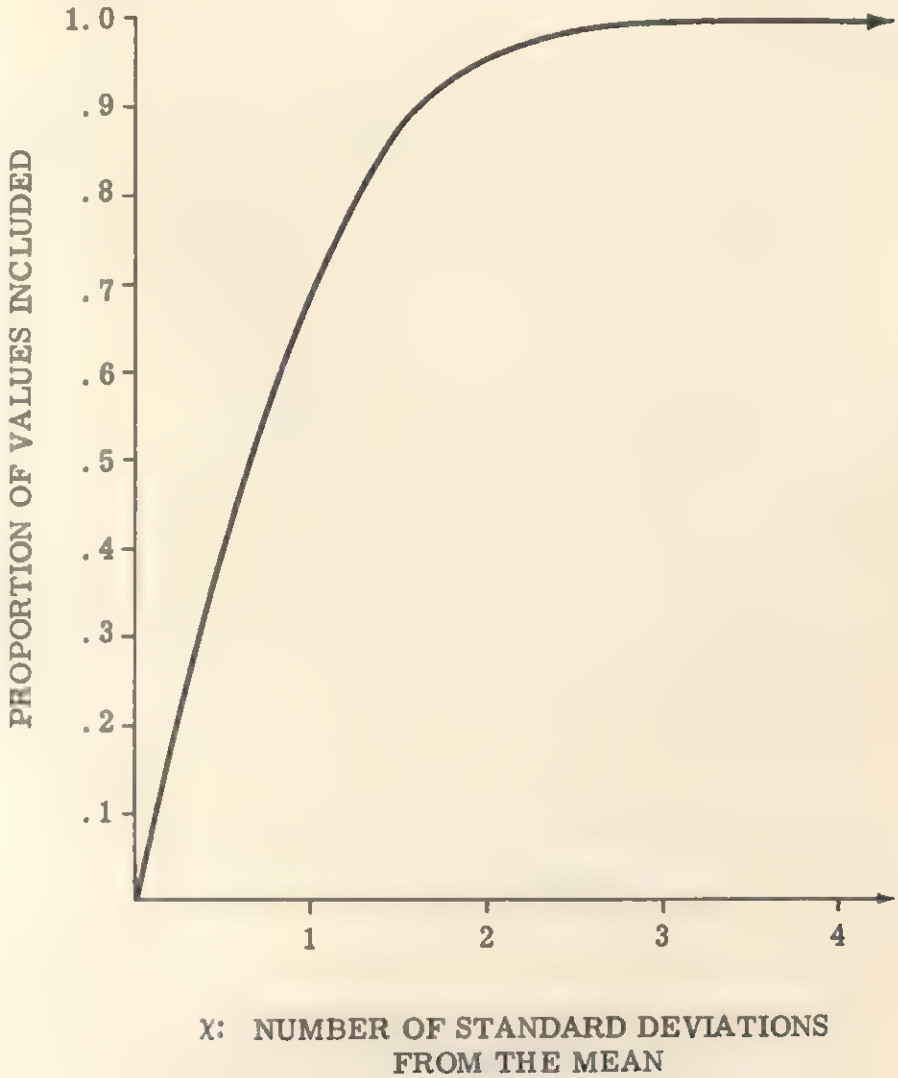
104. Suppose we counted all the "times" of 50 seconds or faster. We would find that exactly five of the observed times fell into this group of times. This group of times would consist of the 32 seconds observed for Subject 7, the 16 seconds observed for Subject 9, the 20 seconds observed for Subject 11, the 15 seconds observed for Subject 18, and the _______ seconds observed for 36
Subject _____. 19

105. We could say the frequency of "times" between zero and 50 seconds was five, since there are exactly five observed times that were less than or equal to _____ 50
seconds.

211. (Continued)

We indicated how a normal distribution could be described by the following graph:

PROPORTION OF VALUES WITHIN χ STANDARD DEVIATIONS OF THE MEAN IN A NORMAL DISTRIBUTION



106. Suppose we counted the "times" between 51 seconds and 100 seconds (including times of 51 seconds or 100 seconds). The time recorded for Subject 1 ______________ (would/would not) be counted, since this time is between 51 seconds and 100 seconds. would

There are exactly ________ observed times that fall between 51 and 100 seconds in the previous table of data. The next group of times that we will consider are those times between 101 and 150 seconds, including any that might be 101 or 150 seconds. Subject 2's time of 128 seconds ____________ (does/does not) belong to the group of times between 101 and 150 seconds. The other times which fall into this group are the time of 149 seconds recorded for Subject 8, the time of 110 seconds recorded for Subject 10, the time of ______ seconds recorded for Subject _____, and the time of 116 seconds recorded for Subject 16.

10

does

120

14

107. We could summarize these frequencies in the following frequency table of **grouped** data.

TIME	FREQUENCY
0-50 sec.	5
51-100 sec.	10
101-150 sec.	5

Notice the three rows correspond to the three different **groups** of times.

208. In the next section, we will consider the advantage of knowing how often different types of samples can be expected to occur when you draw random samples from a particular type of population. It is only because we use _______________ sampling procedures rather than biased sampling procedures that we can use probability theory to calculate theoretical sampling distributions. random

209. For example, suppose we had only asked a group of seniors in the French department for their opinions concerning the overseas campus. This would represent a _______________ sample from the population biased
biased / random
consisting of the 10, 000 opinions of the total student body.

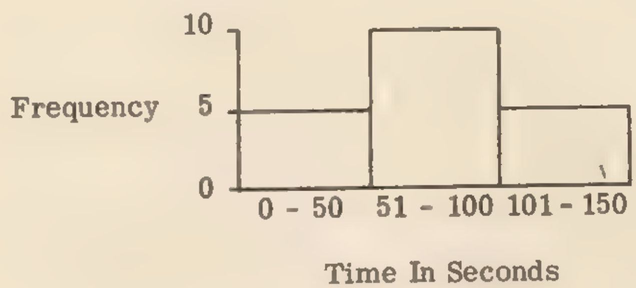
210. With a _______________ sampling procedure, it is possible random
random/biased
to predict the proportion of favorable opinions in the population. If we obtained samples in a _______________ biased
biased/random
manner, it would **not** be possible to calculate a theoretical sampling distribution using probability theory, since our sampling procedure would not be a random process.

211. **Binomial distributions** are only one of several common types of distributions that can be used as a theoretical sampling distribution. The **normal distribution** we considered in an earlier section is another type of distribution often used as a theoretical sampling distribution. We described a normal distribution in terms of the proportion of the values that were within each possible number of standard deviations from its mean.

107. (Continued)

For instance, in the first row "0-50 seconds" indicates that we counted all the times between 0 and 50 seconds. The numeral 5 in the same row but in the frequency column indicates that the ______________ of observed times falling within this group was _________. — frequency (number); five

108. A frequency table of this sort is called a frequency table of **grouped** data because we have determined the frequency for ______________ of values, rather than the frequency for particular values. Below is a graph of the previous frequency table of grouped data: — groups



This frequency distribution gives a reasonable picture of the distribution of times for the different subjects. It indicates clearly that some subjects had times less than 50 seconds, that about the same number had times greater than 100 seconds, and that the most typical times fell between ________ and _________ seconds. — 51, 100

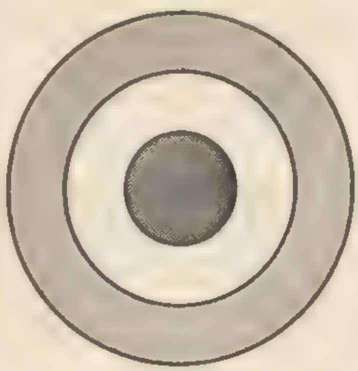
205. Previously, when we considered **experimental** sampling distributions, we were able to determine how frequently to expect different types of samples to occur by actually obtaining a large number of samples. However, it ________ (is/is not) necessary to actually obtain a large number of samples in order to calculate a **theoretical** sampling distribution using the rules of probability. — is not

206. It might seem like magic to be able to obtain a sampling distribution without actually having to collect a large number of samples. When we use probability theory to calculate theoretical sampling distributions, we are simply taking advantage of what has already been determined about random processes. This is perhaps the major advantage of using **random sampling procedures** to collect samples. Probability theory can be used to represent r______________ processes. Random samples are obtained by a ______________ process. Therefore, you can use probability theory to calculate theoretical sampling distributions. — random; random

207. You have seen that the procedure we considered for obtaining samples of students opinions is similar to a simple random process, such as flipping a coin. It is the similarity between the random sampling procedure for obtaining student opinions and the random process of flipping a coin that allows us to use p______________ theory to calculate a theoretical ____________________ distribution for the sample of student opinions without having to actually collect a large number of samples. — probability; sampling

109. If you are interested in obtaining a rough picture of the frequency distribution, it is often useful to group the data in this way — that is, not distinguishing between values that fall within the same group.

Just how much you can group your data and still obtain a sufficiently clear picture of the distribution depends on your particular purposes. For example, in the previous illustration in which subjects were throwing a dart at a target, it was sufficient to divide the board into three circles and assign one of three scores depending upon where each dart struck. Consider the two targets shown below. Target A is divided as the target was in the previous illustration. In Target B, however, more circles have been drawn, thereby dividing the target into narrower rings.



TARGET A

TARGET B

If you assigned a score to each dart depending upon which ring it struck, there would be more possible scores on Target $\underset{A/B}{\underline{\hspace{3em}}}$. The only difference between B

Target A and Target B, however, is that you could distinguish the position of a dart more precisely on Target $\underset{A/B}{\underline{\hspace{3em}}}$. B

200. You could toss a coin seven times and record the number of heads in those seven tosses. You could then repeat this procedure and record the number of heads in those seven tosses. You could repeat this procedure over and over again until you had a very long list of numbers. According to the previous theoretical sampling distribution, you would expect more of these numbers to be ______ (3/2) than you would ______ (3/2). — 3, 2

201. Similarly, according to the previous theoretical sampling distribution, it would be ____________ (more / less) common to obtain a sample of seven tosses in which there were 5 heads than samples of seven tosses in which there were 6 heads. — more

202. Finally, the distribution indicates that samples of seven tosses which were all heads (or all tails) would occur very ______________ (seldom/ often) compared to the other types of samples. — seldom

203. In other words, the types of samples which have a very small probability would be expected to occur very ______________________ (frequently/ infrequently). — infrequently

204. Specifically, if a particular type of sample had a probability of .1, you would expect that type of sample to occur only about _____ time(s) out of every 10 samples obtained in that manner. — 1

110. In a similar way, we might have divided the previous mathematical problem data into smaller groups. For example, we could have divided each of the groups of times into two smaller groups. Instead of considering all the times from 0 to 50 seconds, we would have considered all the times from 0 to 25 and from 26 to 50 seconds. Instead of considering all the times from 51 to 100 seconds as a single group, we would have formed two groups, 51 to 75 and 76 to 100. Finally, we could have divided the group of scores from _______ to _______ seconds into the following two groups: 101 to 125 seconds and 126 to 150 seconds.

101, 150

111. If you become a scientist, one of your chief jobs will be to make observations of variables. Records of these observations are called __________.

data

112. We have seen that one way of presenting data is to arrange it in rows and columns to form a __________.

table

113. A table listing every observed value of the variable is often referred to as a table of _________ data.

raw

114. The number of times a particular value occurs in a table of raw data is referred to as the ____________ of that value.

frequency

115. You __________ (can/cannot) determine the frequency of each value for both a numerical and a non-numerical variable.

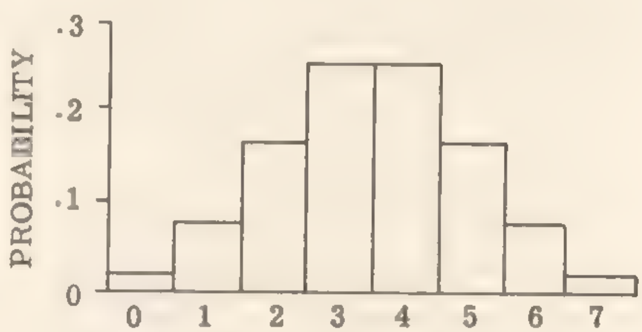
can

197. (Continued)

certain information about the random process, it is possible, using simple rules of probability, to calculate a th____________ sampling distribution. This distribution indicates the relative frequency with which each possible type of sample would be expected to occur if we ____________ the sampling procedure over and over again a very great number of times. — theoretical; repeated

198. Whenever a random process is repeated N times resulting in one of two outcomes each of these N times, you can determine the appropriate b____________ distribution to use as a theoretical sampling distribution if you know the probability of each of the two possible outcomes. — binomial

199. The following ____________ distribution can be used as a theoretical sampling distribution to describe the probability of obtaining any number between 0 and 7 heads out of seven tosses of a fair coin. — binomial



NUMBER OF HEADS IN SEVEN TOSSES OF A FAIR COIN

Notice that the probability of obtaining 3 heads in seven tosses is ____________ (larger/smaller) than is the probability of obtaining two heads in seven tosses. — larger

116. Instead of presenting a complete list of all of the observed values in your data, it is often useful to **summarize** the data in some way. Any term or number summarizing data in this way is called a s__________. statistic

117. One way we can summarize data is to report the frequency of occurrence for each value in the table of raw data. Therefore, each of these frequencies would be a ____________. statistic

118. Since the frequency of a value is determined by counting (enumerating) the number of times it was observed, a frequency is often referred to as an ______________ statistic. enumerative

119. Any value of the variable not observed (one that does not occur in the table of raw data) would have a frequency of _____. If the data consisted of 20 observations, the greatest possible frequency of any value would be _____. 0 20

120. We can summarize the contents of a table of raw data in a table listing the number of times each value occurred in the data. This summarizing table is often referred to as a ______________ table. frequency

121. The enumerative statistics presented in a frequency table indicate how our observations were distributed among the different possible ___________ of the variable we were studying. values

195. (Continued)

the first draw, they were so close to being independent that we could consider them **as if** they were. Therefore, we could treat the two draws as a series of two completions of a random process, where each completion of the process could have only two possible outcomes: "yes" or "no." The theoretical sampling distribution, therefore, was another example of a ________________ distribution. binomial

196. In situations such as this, all that is required to determine the appropriate binomial distribution to use as a theoretical sampling distribution is to determine the probability of each of the two outcomes. Thus, knowing that the probability of heads equals $\frac{1}{2}$ (and that the probability of tails equals $\frac{1}{2}$), each time a fair coin is tossed is sufficient to determine the probability of obtaining each possible number of heads in N tosses of a coin. Similarly, the only information necessary to determine the theoretical sampling distribution indicating the probability of each possible number of "one dot" rolls in N rolls of a die is that the probability of getting "one dot" is ________, whereas the probability of getting "more than one dot" is _________ on each individual roll of the die. $\frac{1}{6}$ $\frac{5}{6}$

197. We will not attempt to consider exactly how a statistician determines the particular binomial distribution to use as a theoretical sampling distribution for any particular number of tosses of a coin or rolls of a die. The procedure is similar (though more complicated) to the procedure we have already used for calculating the theoretical sampling distribution of two tosses of a coin. The important point we wish to make is that, given

122. The **frequency distribution** of a collection of data is the collection of frequencies for the values in that data. Therefore, the enumerative statistics in a frequency table indicate the __________ __________ of the data. — frequency distribution

123. If you were told that a coin had been tossed 100 times and had come up "heads" 75 times and "tails" 25 times, you __________ (would/would not) know the frequency distribution of the data. — would

124. Suppose you tossed a coin 100 times and observed 50 "heads" and 50 "tails." Instead of reporting the actual frequency of "heads" or "tails," you might describe the distribution of the variable by saying that one-half of the tosses were "heads" and that ______ ________ of the tosses were "tails." — one half

125. If you observed 500 "heads" and "500" tails in 1000 tosses of a coin, you could ________ (also/not) say that one half of your observations were heads and one half of your observations were tails. — also

126. When you say one half of the observations are "heads," you are not indicating the actual frequency of "heads" in the data. Instead, you are indicating what **part** of all the observations is "heads." For example, if I told you I tossed a coin some unknown number of times and that one half of the observations were "heads," you __________ (would/would not) know the actual frequency of heads; — would not
you __________ (would/would not) know, however, what part of the data was made up of observations of "heads." — would

193. (Continued)

(for any particular value of N) would be an example of a **binomial distribution**. Thus, the theoretical sampling distribution indicating the probability of obtaining each possible number of heads in 10 tosses of a coin would be an example of a b_______________ distribution. binomial

194. The important feature of repeated tosses of a coin allowing us to represent them with a binomial distribution is that successive tosses of a coin are independent and can be viewed as repeated completions of a simple random process with two possible outcomes. Suppose we grouped the outcomes of rolling a die so as to only consider two possible outcomes: "one dot" or "more than one dot." We could view repeated rolls of the die as a series of completions of a random process with only two possible outcomes ("one dot" or "more than one dot"). A theoretical sampling distribution indicating the probability of obtaining each possible number of one dot outcomes in any series of rolls would be an example of a b_______________ distribution. binomial

195. Statisticians and mathematicians have analyzed processes of this sort so that it is possible for them to determine the particular binomial distribution which would serve as a theoretical sampling distribution for any series of N completions of a simple random process with only two possible outcomes. Earlier, we considered an example in which we obtained a theoretical sampling distribution indicating the probability of obtaining 0, 1, or 2 favorable opinions in a sample of 2 opinions drawn from a population of 10, 000 student opinions. While it is not strictly true that the opinion sampled on the second draw was **independent** of the opinion sampled on

127. If you knew the frequency of a particular value was 10, you would not know whether observations of that value represented a large part of the data or a small part of the data. If a value had a frequency of 10, observations of that value would represent a **larger part** of your data if the data consisted of ______ (12/100) observations rather than ______ (12/100) observations.

12

100

128. Instead of reporting the actual frequencies of each value in your data, it often is useful to indicate the **relative** frequency of each value. The relative frequency of a value indicates how often that value occurred in relation to the total number of observations. For example, if you said the value 10 occurred 30 times in the data, you would be indicating the actual or **absolute** frequency of the value 10. If you said one-third of the total number of observations had the value 10, you would be indicating the ______________ (absolute/relative) frequency of the value 10.

relative

129. Adding together all of the frequencies in a frequency table indicates the total number of observations in the data. To find out what **proportion** or part of the total is represented by a particular frequency, you simply divide that frequency by the total number of observations. If your data consisted of 100 observations and a particular value had a frequency of 50, the proportion of your observations having the value 50 would equal 50 divided by ________, which equals one-half.

100

191. (Continued)

distribution that indicated the probability of obtaining 0, 1 or 2 heads in two tosses of a coin. In much the same manner, we also could have calculated a theoretical sampling distribution that indicated the probability of obtaining 0, 1, 2, or 3 heads out of three tosses of a coin. If we represented the number of times we tossed the coin by the letter N, we know that the fewest heads that could be obtained in N tosses would be 0, whereas the largest number of heads that could be obtained would be N.

In other words, if we tossed the coin seven times, N would equal ______. If we tossed the coin 20 times, N 7
would equal ______. 20

192. You saw how it was possible to calculate a theoretical sampling distribution indicating the probability of obtaining each of the three possible numbers of heads (0, 1, or 2) when we tossed a coin twice (in other words, when N = ____). 2

193. By using much the same procedure, it is possible to calculate a theoretical sampling distribution indicating the probability of obtaining each possible number of heads in N tosses of a coin, no matter what the value of N. For example, we could calculate a ______________ theoretical
sampling distribution indicating the probability of obtaining 0, 1, 2, 3, or 4 heads out of 4 tosses of a coin (when N = ______). The theoretical sampling 4
distribution representing the probability of obtaining each possible number of heads out of N tosses of a coin

130. If your data consisted of only 4 observations and a particular value had a frequency of 1, then the proportion of times that particular value occurred is ______ divided by ________, or one-fourth. 1, 4

131. If your data consisted of 1000 observations and you observed a particular value 250 times, the frequency of that value would represent **one-fourth** of the total number of observations, since 250 divided by 1000 equals __________. one-fourth

132. It is possible to represent a **proportion** in several different ways. For example, you could represent the proportion "one-half" either as a **fraction**, which would be written as $\frac{1}{2}$, or as a **decimal**, which would be written as .5. Similarly, you could represent the proportion "one-tenth" as either a **fraction**, which would be written ________, or as a **decimal**, which would be written __________. $\frac{1}{10}$ .1

If your data consisted of 100 observations of a variable and 10 of those observations had the value 4, you could say the proportion of observations in your data having the value of 4 were ____________ (representing the proportion as a **fraction**) or ____________(representing the proportion as a **decimal**). $\frac{10}{100}=\frac{1}{10}$ .1

133. It is possible to convert a particular fraction into another form representing the same value. For example, $\frac{2}{4}$ is the same as $\frac{1}{2}$. Similarly, you can write a decimal in several ways. For example, .2 is the same as ________. .20

.20/.02

188. (Continued)

For example, while you could not exactly predict the character of any particular sample, you could make the following type of confidence statement on the basis of the previous theoretical sampling distribution:

"I would expect to obtain only about _______ samples in which there were no favorable opinions out of every 100 samples of size 2 obtained in this manner." 9

9 / 20

189. According to the theoretical sampling distribution, the probability (limiting relative frequency) of samples in which no favorable opinions occurred is .09, which could be interpreted as meaning that only about 9 out of every _______ samples would be expected to contain no favorable opinions. 100

190. You have seen how it is possible to represent random processes with probability theory. You have also seen how it is possible to analyze sampling procedures by using probability theory to obtain theoretical sampling distributions. These sampling distributions are useful because they indicate how often you would __________ to obtain each of the possible types of samples. We shall see in the next section how information of this sort can be used to make decisions about populations on the basis of random samples from these populations. expect

191. The theoretical sampling distributions we have just considered are examples of a very important and useful type of theoretical distribution called a **binomial distribution**. We calculated a theoretical sampling

133. (Continued)

Statisticians often find it useful to represent a proportion in terms of **hundredths**. In these terms, one-fourth would equal $\frac{25}{100}$, whereas one-half would equal $\frac{50}{100}$. Similarly, $\frac{3}{4}$ would equal _______ hundredths. 75

134. If a certain proportion equals 25/100, we often say that it equals 25 **percent**. Similarly, if a proportion equals 50/100, we can say it equals fifty **percent**. Representing a proportion as a percentage is simply a convenient way of comparing proportions by converting them all into **hundredths**. For example, you could compare one-half and one-fourth by saying one-half equals __________ 50

50/25

percent, whereas one-fourth equals __________ percent, 25

50/25

(since 1/2 = 50/100 and 1/4 = 25/100.) Similarly, the proportion .25 represents 25/100 or ______ percent. 25

135. If you said the frequency of a particular value in a collection of data represented 50 percent of the data, you would know the frequency of that value represented _______ hundredths of the data. Thus, if the data consisted of 100 observations, a value would have to have a frequency of _______ in order to represent 50 per cent of the observations, since 50 per cent is equivalent to 50/100. A value observed on five out of ten observations also represents 50 percent of those 10 observations since 5/10 equals ______ hundredths. 50 50 50

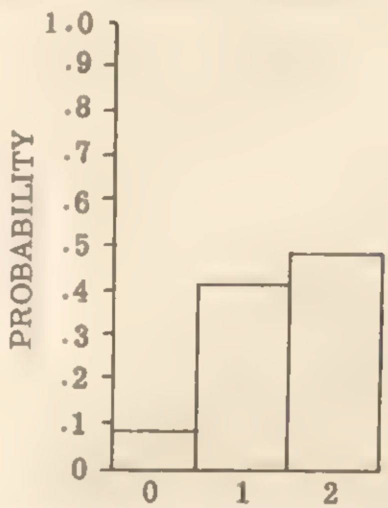
184. The probability of obtaining two **unfavorable** opinions is simply Pr(______), or .09. — Pr(nn)

185. In order to obtain the probability of obtaining one favorable opinion, however, we must apply the **addition** rule and add Pr(______) and Pr(_______). — yn, ny

186. Thus, Pr(one favorable opinion) = Pr(yn) + Pr(ny) = ____ + ____ = ____ (Refer to the previous table for the actual probabilities.) — .21, .21, .42

187. Therefore, by means of a probability analysis, we have determined a ________________ sampling distribution for samples of size two drawn from a population consisting of 10, 000 students' opinions, of which 3, 000 were unfavorable and 7, 000 were favorable. — theoretical

188. The importance of this theoretical sampling distribution (shown on the graph below) is that it tells us what proportion of each possible type of sample to **expect** when we obtain random samples from a population of this sort.



NUMBER OF FAVORABLE OPINIONS
IN A SAMPLE OF SIZE 2

136. If your data consisted of one thousand observations, of which 250 had the value 5, you could say that _______ percent of your observations had the value 5. If 250 out of one thousand observations had the value 5, then 250/1000, or 25/100, of the observations would have the value 5. This is why you could say that 25 p___________ of the observations had the value 5.

25

percent

137. Instead of the word "percent", the symbol % is often written after a number. Thus, 25% means the same thing as 25 _____________.

percent

A proportion written as a percent is often referred to as a **percentage**. Therefore, the following is a list of ____________________s.

percentages

20%, 30%, 5%, 8%, 2%

Remember, ___________ represents a proportion of $\frac{20}{100}$
20%/2%

20%

138. While it is often useful to represent a frequency as a proportion or percentage, it is important to consider the actual frequency. For example, suppose an automobile dealer told you 3/4 of the people to whom he had sold a particular car claimed it was the best car they had ever driven. If he had sold one thousand of these cars, the proportion 3/4 would represent_________ people. On the other hand, if
750/250

he had only sold four such cars, only ______ people would have told him it was the best car they had ever driven.

750

3

Remember, 50% of 1000 is __________, whereas, 50% of 100 is ________.

500

50

181. By using this same reasoning, we could determine the probabilities for each possible pair of opinions. Our results would be those shown in the following table:

Pairs of Opinions	Multiplication Rule	Probability
YY	.7 x .7	.49
YN	.7 x .3	.21
NY	.3 x .7	.21
NN	.3 x .3	.09

In the first column of the table, we have indicated the four possible pairs of opinions. In the second column (in the appropriate row), we have indicated how the multiplication rule would be applied. In the third column, we have indicated the actual probability for each of the possible pairs of opinions. Thus, the probability of obtaining two unfavorable opinions was obtained by multiplying ________ x ________, which indicates .3, .3
indicates Pr(nn) = ________. .09

182. When we considered the samples of two opinions drawn from the population, we were only interested in the number of favorable opinions in the sample. In a sample of size two, it is only possible to obtain 0, ________, or 1
________ favorable opinions. 2

183. We could group the four possible outcomes shown in the previous table, just as we grouped the four possible outcomes for two tosses of a coin where we were only interested in how many heads had occurred. Thus, the probability of obtaining two favorable opinions is simply Pr(yy), or ________. .49

139. It is easy to convert any frequency table into a table listing the proportion of times each value was observed. To do so, you simply divide each frequency in the frequency table by the total number of observations in your data. By doing so, you convert each ______ (absolute/relative) frequency to a(n) ______ (absolute/relative) frequency or proportion. — absolute, relative

140. Consider the frequency table shown below:

VALUE	FREQUENCY	PROPORTION
red	2	
green	3	
blue	5	

The data represented in this table consist of ______ observations of a ______ (numerical/non-numerical) variable. — 10, non-numerical

141. In the third column of the above table, we could list the proportion of observations having each of the three values. For example, the value "red" occurred _____ times in the _______ observations. Therefore, the proportion of times Red occurred equals ____ divided by _______. — 2, 10, 2, 10

142. The proportion of times that the value "green" was observed equals _____ divided by _______, which equals ______. — 3, 10; $\frac{3}{10}$

179. The two draws making up the sample of size two are not **exactly** independent, since the probability of obtaining "yes" on the second draw, given that a "no" was obtained on the first draw, is slightly different from the probability of obtaining a "yes" on the second draw, given a "yes" on the first draw. However, the population is so very large that drawing out a single opinion only alters the probability distribution by a very, very small amount. Thus, we will pretend that the two draws are independent, which allows us to use the ______________ (addition/multiplication) rule to obtain the probability of a "yes" opinion, followed by a second "yes" opinion. multiplication

180. If the probability of obtaining a "yes" opinion is 7/10, then the probability of obtaining two yes opinions in a row (assuming the two draws are independent) is equal to .7 times .7, according to the **multiplication** rule.

$$Pr(yy) = Pr(y) \times Pr(y) = .7 \times .7 = .49$$

All this means is that out of every 100 **pairs** of opinions drawn in this manner, you would expect about 70 of them to begin with a favorable opinion. Of the 70 pairs of opinions that began with a favorable opinion, about 7/10 of them would also have a favorable second opinion. Since 7/10 of 70 is 49, you would expect about 49 of the 100 pairs of opinions to consist of two "yes" opinions. This is indicated by saying Pr(yy) = ________. .49

143. The ______________ of times the value "blue" occurred in the data equals 5 divided by 10, or 1/2. proportion

144.

VALUE	FREQUENCY	PROPORTION
red	2	2/10
green	3	3/10
blue	5	5/10

The numbers in the third column of this table are the ______________ we just calculated. proportions

145. Each of these proportions tells what **part** of the total number of observations is represented by each value. If a particular value had a frequency of 0, you would know that **none** of the observations had that value; therefore, the proportion of observations having that value would equal 0 divided by the total number of observations, which means the proportion would equal ______. 0

146. Suppose you asked 100 people to predict which party, Democratic or Republican, would win the next Presidential election. Your data might look like that summarized in the frequency table shown below:

PARTY	FREQUENCY
DEMOCRAT	75
REPUBLICAN	25

177. (Continued)

represented by NN, NY, YN, and YY, where, for example, NY means a "no" opinion was drawn first and was followed by a ________ opinion. yes

178. As you recall, we were able to determine the probability distribution appropriate for two successive independent tosses of a coin by application of the **multiplication** rule, whereby we multiplied the probability of obtaining a head on the first toss times the probability of obtaining a head on the second toss, to determine the probability of obtaining two successive heads. There is a very slight difference, however, between this situation and the situation involving successive tosses of a coin. As you recall, we stipulated that the multiplication rule applied to successive, **independent** outcomes of a random process. By independent, we meant that the character of one outcome in no way determined the character of the other outcome. Consider very carefully, however, what occurs when two opinions are drawn from a population in the present illustration to form a single sample of size two. If the first opinion in the sample were a "no," we would have removed one of the unfavorable opinions from the population.

Therefore, when the second opinion in the sample is drawn, there will still be 7, 000 favorable opinions in the population but only 2, 999 unfavorable opinions. Thus, the proportion of unfavorable and favorable opinions on the second draw depends upon what opinion was obtained on the first draw. In this sense, the two draws ____________ exactly **independent**. are not

are/are not

146. (Continued)

You could describe your data as consisting of ______ observations of a variable you might call "predicted party," with "Democratic" and "Republican" the two possible __________ of that variable. 100 values

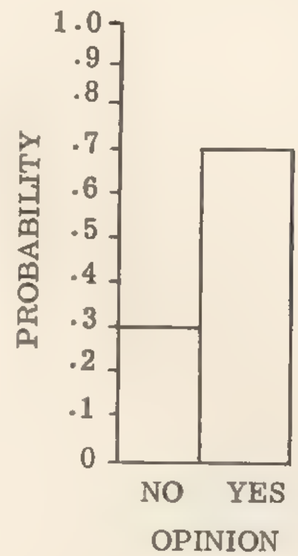
147. The **proportion** of people predicting Democratic ______ whereas the proportion of people predicting Republican was ______. 3/4 1/4

148. The number 75 indicates the ______________ (absolute/relative) frequency of the value "Democratic" in the previous data, whereas the number $\frac{3}{4}$ indicates this value's ______________ (relative/absolute) frequency. absolute relative

149. Just as the frequencies 75 and 25 describe the absolute frequency distribution of the previous data, the proportions 3/4 and 1/4 describe the **relative frequency distribution** of the previous data.

If you tossed a coin 20 times and observed 10 "heads" and 10 "tails," the numbers _________ and _________ indicate the relative (proportional) frequency distribution of your observations. 1/2, 1/2

176. The graph shown below indicates the probability of obtaining a positive opinion and the probability of obtaining a negative opinion when we randomly draw a single opinion from the population. In other words, the graph indicates a ________________ distribution. probability



Notice the similarity between this probability distribution and the probability distribution for a single toss of a coin. In this case, the sample space consists of the two opinions "no" and "yes," whereas, in the case of a toss of a coin, the sample space consisted of the two outcomes "heads" and "tails." With the coin, both possible outcomes ("heads" and "tails") had the same probability, whereas in this case the two outcomes ________ (do/do not) have the same probability. do not

177. Suppose we consider the four possible outcomes that can occur when we draw two opinions from the population. The first opinion we draw can be either a "no" or a "yes." Therefore, the four combinations of opinions could be

150. Suppose you asked 1, 000 people, rather than 100 people, to predict who would win the election. You might have obtained the data represented in the following frequency table:

PARTY	FREQUENCY
DEMOCRATIC	750
REPUBLICAN	250

In the example where we asked 100 people, 3/4 of them predicted "Democratic" and 1/4 predicted "Republican." In the present example of 1, 000 predictions, the proportion of people predicting "Democratic" is _____ 3/4
whereas the proportion of people predicting "Republican" is _______. 1/4

151. The important point to be made here is that even though the _______________ (absolute/ relative) frequency distributions were absolute
different for the two examples, the _______________ (absolute/ relative) relative
frequency distributions were the same.

152. Let us consider another case in which we change an absolute frequency distribution to a relative frequency distribution by converting each frequency to a proportion. Suppose you have the data represented in the following frequency table:

VALUE	FREQUENCY
0-5	3
6-10	5
11-15	2

173. What we are really interested in, however, is not which particular student's opinion is sampled, but the nature of his opinion. Therefore, it would be appropriate to **group** together the opinions of all students who answered **yes** and all students who answered **no**. If 3,000 students answered "no," we would **add** the probabilities of sampling each of their opinions to find the probability of obtaining any opinion from the group of negative opinions. Therefore, since each of the 3,000 negative opinions has a probability of 1/10,000, the probability of obtaining 1 or another of these negative opinions is simply 3,000 divided by 10,000, or ________. 3/10

3/10 - 3/100

174. Notice also that the probability of obtaining a negative opinion is simply the proportion of negative opinions in the population. This is a characteristic of all sample spaces in which every outcome has been assigned the same probability. In other words, if every outcome in a sample space has the same probability and if we combine some number of these outcomes to form a group of outcomes, the probability of sampling any member of the group is simply the number of outcomes in the group divided by the total number of outcomes. For example, if a sample space consisted of ten outcomes, each having the same probability, the probability for obtaining any member of a group of three of these outcomes is equal to ____/10. 3/10

175. Returning to our illustration, we can say that the probability of obtaining a negative opinion when we draw a single opinion from the population is equal to 3/10. Therefore, the probability of obtaining a positive opinion is ________ since there are only two possible outcomes and since the sum of their probabilities must equal 1. 7/10

152. (Continued)

Note that this data consists of ______ observations of a ______ variable. 10 numerical

10/12 — numerical/non-numerical

Furthermore, the frequencies in the table are for ______ data. grouped

grouped/ungrouped

153. The number 3 represents the number of times a value between _____ and _____ occurred in the data. 0, 5

154. The number of times a value between 11 and 15 was observed is _____. 2

155. The group of values between _____ and _____ occurred more often than any other group of values in the data. 6, 10

156. The **proportion** of times a value was observed between 6 and 10 is equal to _____ divided by _____, which equals 1/2. 5, 10

157. The proportion of times a value between 0 and 5 was observed is ________, whereas the proportion of times values between 11 and 15 were observed is _____. 3/10 2/10 (1/5)

158. Therefore, the ______________ frequency distribution of the previous data is represented by the 3, 5, and 2, whereas the ______________ frequency distribution is represented by the numbers 3/10, 5/10, and 2/10. absolute relative

absolute/relative — absolute/relative

170. While the theoretical sampling distribution shown in Graph B is based on a rather simple reasoning process, the same general procedure can be used to determine a theoretical sampling distribution for much more complicated random sampling procedures. As an illustration of how we can use the same reasoning procedure to determine theoretical sampling distributions in other situations, let's consider how we might determine the theoretical sampling distribution for the same situation for which we derived an experimental sampling distribution. As you recall, we considered a situation in which random samples of size 2 were obtained from a population of 10, 000 student opinions. Let's suppose that 3, 000 of the student opinions were negative (the students answered "no"). Therefore, __________ of the students' opinions were positive (the students answered "yes"). 7, 000

171. By thoroughly mixing all of the student opinions together in a large basket, we hoped to give each student's opinion an equal opportunity of appearing in the sample. In other words, we hoped that this was a ______________ sampling procedure. random

172. Since there were 10, 000 student opinions, each opinion could be regarded as having the same probability because, after many repeated samples of this sort, we would expect each opinion to be sampled about the same proportion of times. If we viewed the 10, 000 student opinions as a sample space consisting of 10, 000 possible outcomes, each outcome could be assigned a probability 1/10, 000 with the sum of the probabilities equal to ____ 1

159.

TABLE A

VALUE	FREQUENCY
0 - 5	5
6 - 10	3
11 - 15	2

TABLE B

VALUE	FREQUENCY
0 - 5	300
6 - 10	500
11 - 15	200

Of the two tables shown above, the table having the same **relative** (proportional) frequency distribution as the data shown in Frame 152 is Table ________. B
A/ B

160. Notice that Table B consists of data from ________ observations. Thus, each frequency would be converted to a proportion by dividing it by ______. 1000 1000

161. Expressing the frequency of a value as a proportion indicates what part of the total number of observations were of that particular value. The largest possible proportion you could obtain would occur when all the observations in your data had the same value. This proportion would be ________. one
For example, if the frequency of a particular value were 1, 000 in a collection of 1, 000 observations, the relative frequency (proportion) of observations having that value equals ________ divided by ________, or 1. 1, 000 1, 000

162. If a particular value had a frequency of 0, it would be represented in a relative frequency distribution by a proportion of ________. 0

168. Earlier, we considered an illustration in which we obtained an **experimental** sampling distribution for samples of size 2 obtained from a population of 10, 000 student opinions (where each opinion was either for or against a particular proposal). It is possible to view each pair of tosses of a coin as a sample of size 2 from an infinite population consisting of an umlimited number of tosses. In that sense, the probability distribution in Graph B is a sampling distribution for samples of size 2 from an unlimited population. In this case, however, we have determined the sampling distribution **not** by actually obtaining samples, but through a reasoning process using probability theory. Thus, the sampling distribution shown in Graph B can be thought of as a(n) ______________ sampling distribution. theoretical
experimental/theoretical

169. To obtain this **theoretical** sampling distribution, we begin with a careful consideration of what we mean by a fair coin. Then, by carefully reasoning out the implications of our definition of a fair coin, we arrived at the conclusion that the sampling distribution shown in Graph B should represent the sampling distribution for pairs of tosses of a coin. We have also seen how probability theory can be used to represent a random process, such as tosses of a coin, and how simple rules, such as the **addition** and **multiplication** rule, can be used to represent characteristics of this random process. The advantage of random processes by probability theory is that (if our representation is appropriate) it is possible to use the rules of probability theory in a mechanical fashion, to arrive at a ______________ theoretical
experimental/theoretical
sampling distribution.

163. Notice that all of the proportions in a relative frequency distribution must lie somewhere between the largest possible proportion of ______ and the smallest possible proportion of ______. 1 0

164. Each of the proportions in a relative frequency distribution represents the part of the total collection of observations having a particular value. All of the parts of something added together must equal the whole. Therefore, the sum of all the proportions in a proportional distribution must always equal ______. 1

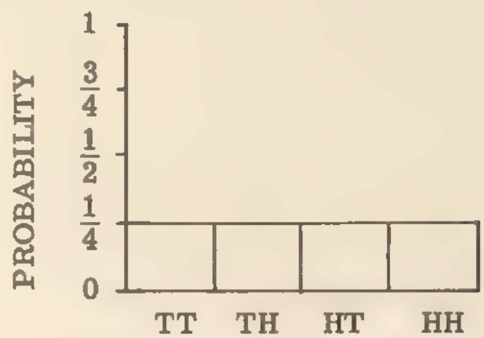
165. A little thought will indicate why the total of all of the proportions in a proportional distribution must equal 1. Each proportion is a fraction whose numerator is the frequency of some value and whose denominator is the total number of observations represented in the data. The sum of all these fractions would simply equal the sum of all the numerators over this common denominator. However, the sum of all the numerators is the sum of the frequencies in the frequency table, which ____________ (does/does not) equal the total number of observations. Therefore, the sum of all of the fractions or proportions would always equal ________. does 1

For example, $\frac{2}{10} + \frac{5}{10} + \frac{3}{10}$ equals $\frac{2 + 5 + 3}{10}$, which equals $\frac{10}{10}$, or ______. 1

166. We stated earlier that a statistic is any number or term that summarizes a collection of data. Each of the proportions in a proportional distribution summarizes something about the frequency of a particular value in the collection of data. Therefore, each proportion ______________ (would/would not) be called a statistic. would

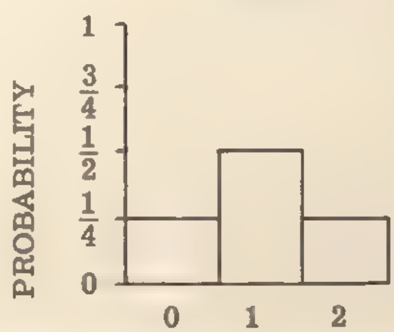
167. On Graph A (shown below) we have indicated the probability distribution on a sample space consisting of 4 possible outcomes of 2 tosses of a coin. Graph B represents the sampling distribution on a sample space consisting of 3 possible outcomes for 2 tosses of a coin. Notice that in Graph B we only consider the number of heads occurring in each pair of tosses. Getting 0 heads in two tosses is equivalent to the outcome TT on Graph A. (Similarly, 2 heads on Graph B is the same as the outcome HH on Graph A.) So the probability of this event is the same on both graphs. However, obtaining "1 head" on Graph B could be either the outcome TH **or** the outcome HT, shown in Graph A. According to the simple addition rule of probability theory, therefore, the probability of obtaining 1 head is equal to the probability of TH **plus** the probability HT (i.e., the probability of obtaining 1 head equals ________.

$\frac{1}{2}$



OUTCOMES OF TWO TOSSES

GRAPH A



NUMBER OF HEADS IN TWO TOSSES

GRAPH B

167. Of the three tables shown below, two are

_______________ frequency tables and the other is a — absolute
absolute/relative

_______________ frequency table. — relative
absolute/relative

TABLE A

TIME	FREQUENCY
0-100 sec.	10
101-200 sec.	90

TABLE B

TIME	FREQUENCY
0-100 sec.	100
101-200 sec.	900

TABLE C

TIME	FREQUENCY
0-100 sec.	1/10
101-200 sec.	9/10

168. Table C is a relative frequency distribution describing

the data in Table _______________. — A or B
A/B/A or B

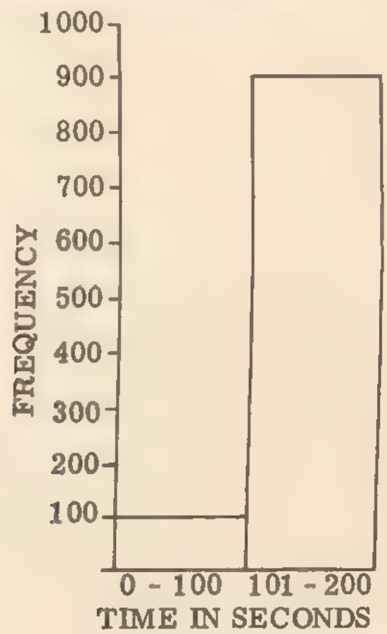
169. Even though the enumerative statistics in Table A are different from those in Table B, both of these absolute frequency distributions can be represented by the same _______________ frequency distribution. — relative (proportional)

163.	Similarly, the outcome ______ (HH/HT) would represent 2 heads in the new list of 3 outcomes.	*HH*
164.	On the other hand, **both** the outcome ______ and the outcome ______ shown in Graph B correspond to the outcome one head in the new list of 3 outcomes.	*HT* TH
165.	In other words, in the new sample space of 3 possible outcomes, we **group** the outcomes HT and TH, since they both represent the occurrence of one head. According to the simple **addition** rule, the probability of obtaining either HT **or** TH is equal to the ________ of their individual probabilities.	sum (total)
166.	We could state this formally by stating Pr(1 head) equals Pr(HT) + Pr(TH). Both Pr(HT) and Pr(TH) are equal to $\frac{1}{4}$, as indicated on Graph B. We could substitute the actual values of these probabilities into the previous equation, which would give us Pr(1 head) $= \frac{1}{4} + \frac{1}{4} =$ ______	$\frac{1}{2}$
	According to this reasoning, in other words, the probability of obtaining no heads is simply the probability of obtaining a tail on the first toss and a tail on the second toss (which is equal to $\frac{1}{4}$). The probability of obtaining 2 heads is simply equal to the probability of a head on the first toss and a head on the second toss (which is equal to $\frac{1}{4}$). To find the probability of obtaining one head, however, we must **group** the two outcomes shown previously on Graph B, to form a new outcome entitled "one head," whose probability equals the **sum** of the probabilities of the two individual outcomes in the group. Therefore, the probability of one head equals ________.	$\frac{1}{2}$

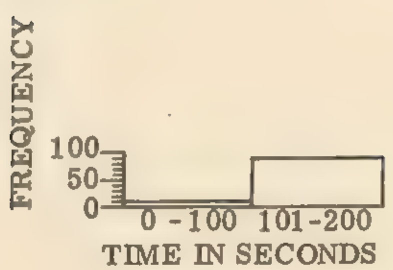
170. Just as you could present an absolute frequency distribution on a graph, it is also possible to present a proportional distribution in graphic form. In the distribution graph, we represented the frequency of each value by the __________ of a column. In this same way, we may represent each proportion in a proportional distribution. height

171. The highest possible column that could ever occur on a proportional distribution graph would represent a proportion of _________. The three graphs shown below represent the same data shown just previously in tables A, B, and C. Notice that Graph C is a __________________ frequency distribution graph, just as Table C was a proportional frequency table. 1 proportional

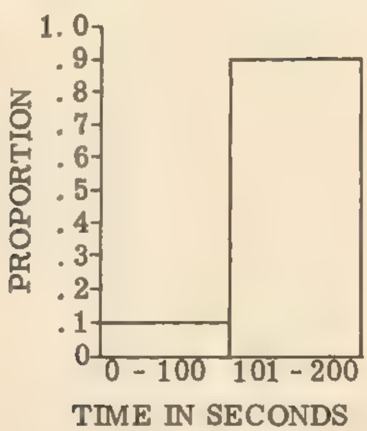
GRAPH B



GRAPH A

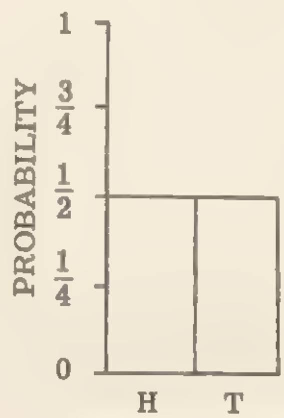


GRAPH C



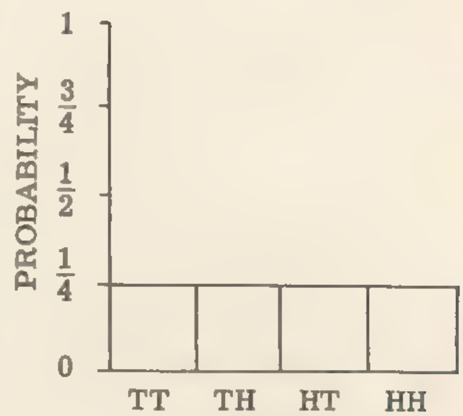
161. You also considered an illustration in which we determined the probability distribution for a sample space consisting of the 4 possible outcomes for two tosses of a coin, using the simple multiplication rule of probability theory. We determined the probability distribution shown in Graph ______ (below) by noting B
A / B

that the probability of each of the 4 possible outcomes in that sample space was equivalent to the product of 2 probabilities shown in the other distribution.



OUTCOME OF
ONE TOSS

GRAPH A



OUTCOME OF
TWO TOSSES

GRAPH B

162. Let us go one step further in our analysis of the random process of tossing a coin. Suppose we **grouped** the 4 possible outcomes of a pair of coin tosses on the basis of how many heads had occurred. Describing the pairs of tosses solely in terms of the number of heads that have occurred would define 3 possible outcomes: no head, or two heads. In other words, the outcome represented by TT in the Graph B (above) would correspond to the outcome ______ heads in the new list no
of 3 possible outcomes.

172. Notice that it is difficult to compare the two absolute frequency distributions in Graph A and B because the number of observations in each collection of data was

______________.
the same/different

different

173. The difference in the size of the collection of data could easily obscure the similarity of the two distributions. This similarity is the fact that the ______________
absolute/relative
frequency distributions of the two groups of time are the same.

relative

174. Therefore, if you wish to compare two collections of data when there are different numbers of observations in each collection, it is often useful to use a(n)

______________ or ______________
absolute/relative
frequency.

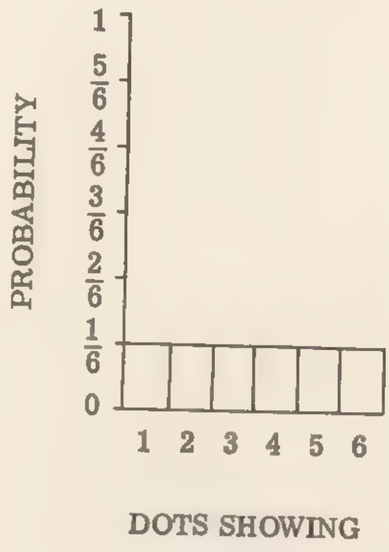
relative, proportional

159. Let's review the two rules we have just considered:

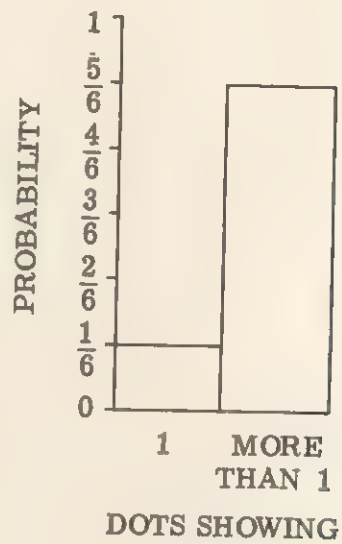
1) the simple **addition rule**
2) the simple **multiplication rule**

By using the ________________ rule, we were able addition
(addition/ multiplication)

to determine the probability distribution for the sample space shown in the Distribution B (below), since it is based on grouping some of the outcomes in the sample space shown in Distribution A.



DISTRIBUTION A



DISTRIBUTION B

160. In other words, you determine the probability of the group of outcomes represented by the title (more than one) by adding the probabilities of the individual outcomes in the group. This, therefore, was an application of the ________________ rule. addition

REVIEW I

FILL IN THE BLANKS:

1. If a value is not recorded in the data, the value has a frequency of __________. zero

2. A collection of frequencies in a frequency table is called a frequency _____________. distribution

3. Let us assume that we have two red marbles, five green marbles, and four blue marbles in a box. The proportion of red marbles is __________. $\frac{2}{11}$

4. Let us assume that we have two groups in an experiment. Group A has ten brunettes and five blonds. Group B has twenty brunettes and ten blonds. The relative frequency of the two groups is ________________. the same
   the same/different

5. The largest possible proportion is _________. 1

6. The sum of the proportions in a proportional distribution must equal __________. 1

7.

Totals				
Student	1	2	3	4
Bob	4	0	9	14
Mary	2	1	8	18
Lon	3	7	12	19

What is the score in the second row, fourth column?

_______ 18

155. We just illustrated why the probability of a head on one toss followed by a head on a second toss should equal $\frac{1}{4}$. The simple **multiplication rule** of the probability theory states that the probability of obtaining a particular sequence of two independent outcomes is equal to the **product** of the probabilities of each outcome. We can state this in terms of the previous illustration as

$$Pr(HH) = Pr(H) \times Pr(H).$$

Since we know $Pr(H) = \frac{1}{2}$, we can substitute actual values to obtain,

$$Pr(HH) = \frac{1}{2} \times \frac{1}{2} = \_\_\_\_\_\_$$ $\frac{1}{4}$

156. This formula can be interpreted verbally as follows: the probability of obtaining a head on one toss and a head on the next toss is equal to the probability of obtaining a head on the first toss ____________ the probability of obtaining a head on the second toss. times(multiplied by)

157. The **multiplication rule** is simply a formal way of representing a characteristic of the process we have already considered. In other words, you would expect about $\frac{1}{2}$ of the pairs of tosses of a fair coin to begin with a head. Of the $\frac{1}{2}$ of the pairs that began with a head, you would expect $\frac{1}{2}$ of them also to be a head on the second toss. The m______________ rule simply states that of the pairs beginning with a head (about $\frac{1}{2}$ of all pairs), $\frac{1}{2}$ of these would end with a head. Thus, $\frac{1}{2}$ of $\frac{1}{2}$ is equal to $\frac{1}{2}$ **times** $\frac{1}{2}$, or __________. multiplication $\frac{1}{4}$

158. Similarly, you would apply the multiplication rule to find the probability of a head followed by a tail, as follows:

$$Pr(HT) = Pr(\_\_\_) \times Pr(\_\_\_)$$ H, T

MULTIPLE CHOICE:

8. A list of things, one underneath the other, is called a:

   a. column.

   b. row.

   c. radial.

   d. none of the above ______ a

9. A list of things arranged side by side is called a:

   a. column.

   b. radial.

   c. row.

   d. none of the above ______ c

10. If we list all the observations and do not summarize them, we have:

    a. finished data.

    b. theoretical distribution.

    c. raw data.

    d. none of the above ______ c

11. A number, or term, that summarizes or describes a collection of data is called:

    a. raw data.

    b. a statistic.

    c. a theoretical distribution.

    d. none of the above ______ b

150. It would seem reasonable to assign each of these four outcomes the probability ________, since each outcome would then have the same probability, and the sum of the outcomes would equal one. — 1/4

151. Notice that Pr(HT) = 1/4 can be interpreted to mean that about one out of every four, or about 25 out of every 100, or about ________ out of every 1, 000, pairs of tosses will begin with a head and end with a tail. — 250

152. Pr(HH) = $\frac{1}{4}$ **does not** mean that exactly 1 out of every 4 (or 25 out of every 100) pairs of tosses will consist of a head on the first toss and a head on the second toss.

Pr(HH) = $\frac{1}{4}$ means that, in the long run, about _____ out of every 4, or ________ out of every 100, pairs of tosses will consist of a head on the first toss and a head on the second toss. — 1, 25

153. The **multiplication rule** of probability theory is a formal way of representing this relationship between **independent** outcomes of a random process. We stated earlier that if two outcomes were **independent**, the character of one outcome has no bearing on the character of the second outcome. For example, the outcome of one toss of a coin has no bearing on (does not determine) the outcome of another toss of the coin. Therefore, the two tosses are said to be ind____________. — independent

154. Similarly, when you roll a die, the outcome of one roll in no way determines the outcome of another roll. Therefore, rolls of a die are said to be ____________. — independent

12. Frequencies are often called:

a. raw data.

b. inoperative.

c. enumerative statistics.

d. none of the above ______ c

TRUE OR FALSE:

13. We refer to something that does not change during an experiment as a constant. ______ true

14. Something that does change in an experiment is called a variable. ______ true

15. The numbers one to ten are written on pieces of paper and placed in a hat. The number 4 is drawn from the hat. The number 4 is an observed value. ______ true

16. Let us assume that we have an ordinary die. On such a die, the number 7 is a possible value. ______ false

17. The number of nickels in a cookie jar would be an example of a continuous variable. ______ false

18. Numerals can only be used to represent numbers. ______ false

19. Only a list of observed values can be referred to as data. ______ true

144. You could use precisely the same reasoning to determine approximately how many pairs would begin with a head and end with a tail. Thus, of the 100 pairs, you would expect about _______ of them to begin with a head. Of these 50 which began with a head, you would expect about _______ of them to end with a tail.

50

25

145. Thus, of 100 pairs of tosses, you would expect about ________ pairs to begin with a head and end with a tail.

25

146. It is clear that you would expect about 25 of each of the four possible pairs: HH, HT, TH, TT, where, for example, HT represents a head on the first toss and a tail on the second toss, and TH represents a __________ on the first toss and a __________ on the second toss.

tail

head

147. Imagine a **sample space** consisting of the four possible outcomes for a pair of tosses: HH, _____, TH, TT.

HT

148. Note that this is a perfectly acceptable sample space, since **one and only one** of these outcomes can occur with each **pair** of tosses. In other words, the four outcomes are m________________ e__________________ and e__________________.

mutually exclusive

exhaustive

149. You would expect each of these four possible outcomes to occur about the same number of times. In other words, you would expect each of these possible outcomes to occur about _______ times out of every 100 pairs of tosses.

25

The review that you have just completed can assist you in evaluating your progress at this point in the program. If you had no difficulty with the review, proceed to the next section. If you did have trouble with any of the review questions, return to the place in the program where this material is presented and make sure you understand the material before going on to the next section.

Follow this procedure with each of the reviews in the program.

139. Let's consider only those pairs of tosses beginning with the toss of a head. In other words, we know the first toss in each one of these pairs of tosses is a ________. head

140. Since a coin cannot **remember** what happened to it on the previous toss, the second toss in each pair of tosses is said to be **independent** of what occurred on the first toss. In other words, the behavior of a coin on the second of two tosses ____________ (does/does not) depend upon how the coin fell on the first toss. does not

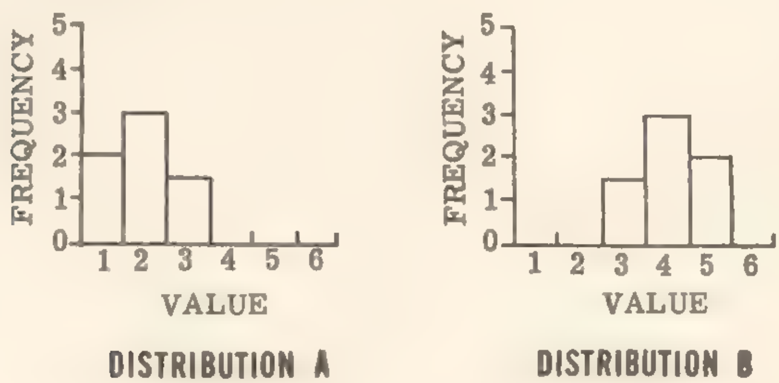
141. Successive tosses of a coin are said to be ind_________, since the outcome of any one toss **does not depend** upon what has occurred on any of the other tosses. independent

142. Therefore, considering only the 50 pairs of tosses which began with the toss of a head, you would expect about _______ of the second tosses in each of these 50 pairs to be a head since about 1/2 of a series of tosses should be heads. 25

143. To summarize, out of 100 **pairs** of tosses of a fair coin, you would expect about 50 (1/2 of the 100) to begin with a head. Of these 50 pairs that **began with a head**, you would expect about 25 (1/2 of the 50 pairs) to end with the toss of a tail. In other words, out of the 100 pairs of tosses, you would expect about ______ of them to consist of a head followed by a head. 25

1. You have seen how any collection of data can be regarded as a distribution of values. By "distribution" we mean the number of times each of the possible values has been ________________. — recorded (observed)

2. Next, we shall consider some of the ways in which distributions differ and how these differences can be described. Consider, for example, the two distributions shown below.



You would probably say that the typical observed value in Distribution B is ________________ (larger/ smaller) than the typical value of Distribution A. — larger

3. While the two distributions have approximately the same shape on the graph, the **center** of Distribution B is at a ________________ (larger/ smaller) value than the center of Distribution A. — larger

137. Suppose you were to toss a coin twice. You could list four mutually exclusive and exhaustive outcomes for this process as follows.

Outcome 1:

a **head** on the first toss and a **head** on the second toss.

Outcome 2:

a **head** on the first toss and a **tail** on the second toss.

Outcome 3:

a **tail** on the first toss and a **head** on the second toss.

Outcome 4:

a **tail** on the first toss and a ______ on the second toss. tail

138. If you were using a **fair** coin, the probability of heads would equal the probability of tails on each toss of the coin. In other words, Pr(heads) = Pr(tails) = ______. $\frac{1}{2}$

Bearing this in mind, let's consider how you would assign probability values to a sample space consisting of the four possible outcomes of the **two** tosses of a coin that were listed above. First of all, let's consider about how many heads you would expect to obtain in 100 tosses of the coin. If the coin were fair, you would expect about ______ heads to occur in 100 tosses. 50

Therefore, if you tossed a coin twice and repeated this procedure 100 times (if you made 100 **pairs** of tosses), you would expect about ______ of the 100 pairs to begin with the toss of a head. 50

4. You could say the **central tendency** of Distribution A is different from the central tendency of Distribution B. In other words, if the values in one collection of data are generally larger than the values in another collection of data, you could say the distributions of the two collections of data have ________________ **central** tendency. — a different
the same/a different

tendency.

5. Consider the three collections of data listed below:

Data A: 2, 4, 3, 12, 6, 4

Data B: 10, 12, 14, 16, 50

Data C: 1, 3, 2, 3, 2, 6

Notice that the values in Data A are **more** similar to the values in Data ______ than they are to the values in — C
C/B

Data ______. — B
C/B

6. The distributions of Data A and Data B seem to have ______________ central tendencies, whereas the — different
similar/different

distributions of Data A and Data C have ______________ — similar
similar/different

central tendencies.

7. If the values in two distributions were quite similar, we would describe the two distributions as having similar

__________ ______________. — central tendencies

8. Consider the two sets of data shown below.

Data A: 4, 3, 6, 6, 4, 6

Data B: 21, 23, 20, 21, 23, 20

You could describe Data ______ as having a larger typical — B
A/B

value than the other collection of data.

133. The second outcome in Distribution B (more than one dot showing) consists of five of the possible outcomes shown in Distribution A. Applying the simple a________ rule of probability theory, we have assigned the second outcome in Distribution B, a probability equal to the total of the probabilities assigned to the five individual outcomes included in that **group** of outcomes. — addition

134. The **addition rule** as it would apply to outcomes in the type of sample space we have considered can be stated formally in the following manner: If A and B are two different outcomes in a sample space, the probability of either outcome A **or** outcome B is simply the probability of A **plus** the probability of B. Therefore, we could represent the **addition** rule symbolically as:

Pr(A or B) = (Pr(___) + Pr(___) — A, B

135. Extending this rule, we would apply it to the previous illustration as follows.

Pr(more than one dot) = Pr(two dots or three dots or four dots or five dots or six dots) =
Pr(two dots) + Pr(three dots) + Pr(four dots) +
Pr(five dots) + Pr(six dots)

Substituting the actual probability values, we could write

$$\text{Pr(more than one dot)} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \_\_\_\_$$ — $\frac{5}{6}$

136. To summarize, the simple __________ rule of Probability Theory says that when we group outcomes in a sample space, the probability of an outcome in the group occurring is simply the _______ of the probabilities of each of the outcomes in the group. — addition; sum (total)

8. (Continued)

Suppose the two collections of data represented the ages of people at two different parties. In other words,

Data _____ would have been collected at a children' s A
A/B

party and Data _____ would have been collected at a B
A/B

party for young adults.

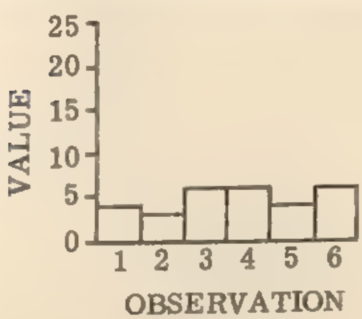
9. A person who was 5 years old would be more representative of people at a party corresponding to Data _____ than A
A/B

he would be to the data collected at the other party.

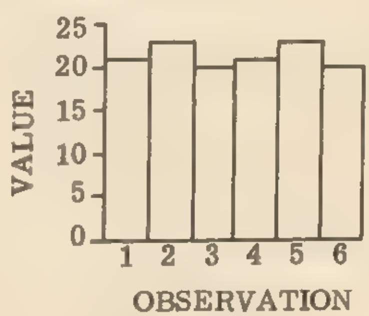
10. Even though it does not occur in Data A, the number 5 is more representative or typical of that collection of data than it is of the other collection. In a similar sense, the number 22 is more typical of Data _____ than it is of B
A/B

the other collection of data.

11. The two previous collections of data differ mainly in the magnitude or size of the recorded values. This can be illustrated by the two graphs of raw data shown below.



DATA A



DATA B

Notice how the values in Data _____ all seem to be close B
A/B

to (cluster around) the value 22, whereas the values in the other data seem to cluster around the value _____. 5
5/10

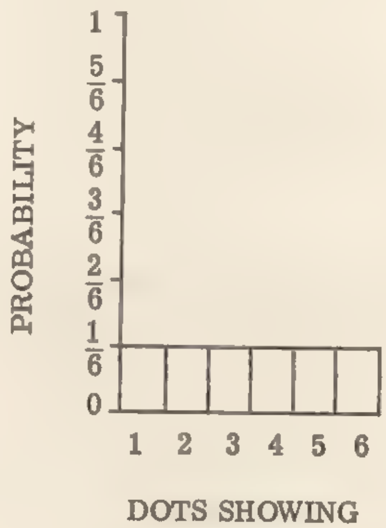
131. The probability of obtaining more than one dot is simply the **sum** of the probabilities for each of the five outcomes grouped in this category. This could be stated symbolically as follows:

Pr(more than one dot) = Pr(two dots) + Pr(three dots) + Pr(four dots) + Pr(__________ dots) + Pr(__________ dots).

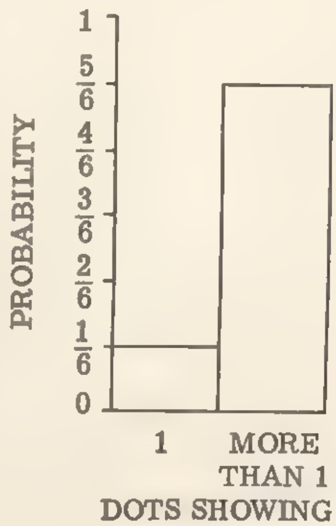
five

six

132. This application of the simple **addition rule** can also be illustrated by the following two graphs:



DISTRIBUTION A



DISTRIBUTION B

Note that Distribution A is the ____________________ distribution on the sample space consisting of six possible outcomes, whereas Distribution B is the ______________ ________________ on the sample space consisting of ______ possible outcomes.

probability

probability
distribution
two

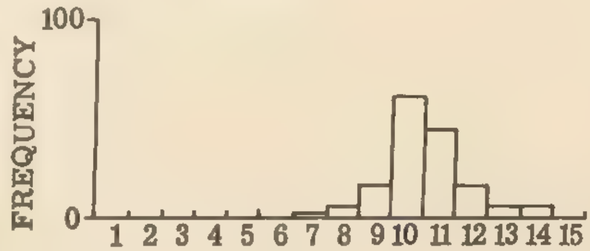
12. Suppose a student named John received grades of 75, 70, 76, 75, 76 in a course. Another student, Dick, received grades of 95, 95, 96, 75, 98. A grade of 75 would be more typical or representative of ______________ (John's/Dick's) grades, even though both students had received a grade of 75 during the course. John's

13. The reason a grade of 75 is more typical of John's grades is that a score of 75 was unusual for Dick. Most of Dick's grades seem to be clustered around ________ (95/72). 95

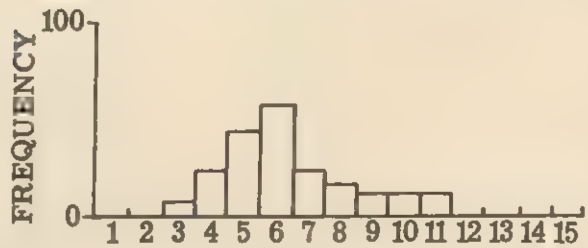
We would say, therefore, that the central tendency of John's grades seems to be ____________ (higher/lower) than the central tendency of Dick's grades. lower

14. If someone asked you to characterize John's work by reporting his typical grade, you would be more likely to answer ______ (75/85) than you would ______ (75/85). 75, 85

15. DISTRIBUTION A



DISTRIBUTION B



Notice how the values in Distribution A seem to be clustered around the value ______ (5/10), whereas the values 10
in Distribution B are clustered around the value ______ (5/10): 5

126. The phrase **expected frequency** in the heading of column two implies each of the six possible outcomes will not necessarily occur exactly 100 times out of the 600 rolls. "**Expected frequency**" implies that each of the six possible outcomes would occur **about** equally often "in the long run" (in a very long series of successive rolls). Thus, ________ is our best bet concerning the number of times we would expect each of these six outcomes to occur in a series of 600 successive rolls. 100

127. In Table B, we have indicated the two possible outcomes in the new sample space. Naturally, we would still expect about ________ rolls of a single dot in a series of 600 successive rolls of the die. 100

128. The second outcome shown in Table B represents all rolls in which "**more than one dot**" was shown on the face of the die. **Expected frequency** of all outcomes falling in this group is simply the ___________ of the frequencies for the five outcomes that fall into this group. sum (total)

129. According to Table B, the expected frequency of a single dot in 600 rolls of a die is ________. Therefore, the expected frequency of "more than one dot," which is the only other possible outcome, is ________. 100 500

130. If a particular outcome is expected about 500 times out of 600 rolls, it would have a probability of $\frac{500}{600}$, or ______-sixths. 5

16. In other words, you could describe the difference between the two distributions shown in the previous frame by pointing out that the **central tendency** of one distribution was different from that of the other. For example, if you said you were referring to the distribution whose typical or central value was 10, it would be clear that you were referring to Distribution ______ (A/B) rather than to the other distribution. A

17. The characteristic of distributions we have been referring to as the "typical value" is generally a value near the "center" of the distribution. In other words, we can usually think of the values in a distribution as being clustered around (close to) a typical or central value. That is why we refer to this characteristic of a distribution as its central __________. tendency

18. Up to this point we have not been very specific about what we mean by "central tendency." We have purposely not done so because there is more than one way to define the typical value or central tendency of a distribution. In other words, there ______ (is/is not) more than one acceptable way of defining the central tendency of a distribution. is

19. If one value occurred more frequently in a distribution than any other value, you might consider that value the most common or typical value of the distribution. Therefore, one way of characterizing the central tendency of a distribution ________ (would/would not) be to report the most frequently occurring value in that distribution. would

123. (Continued)

six possible outcomes. These five outcomes have been grouped under the heading "more than one dot." These five outcomes are: two dots, three dots, ________ dots, four
____________ dots, and __________ dots. five, six

124. According to the probability distribution on the sample space consisting of six outcomes, each of the six possible outcomes would be expected to occur about at the same frequency in any long series of rolls. Thus, if we rolled the die 600 times, we would expect each of the six possible outcomes to occur approximately ______ 100
times.

125. In the new sample space we **group** all outcomes in which more than one dot are showing. (We do not distinguish between outcomes in which more than one dot is showing.) The tables shown below indicate the expected frequency of each of the possible outcomes in 600 rolls of the die in the two sample spaces.

Dots showing	Expected Frequency in 600 rolls	Dots showing	Expected Frequency in 600 rolls
one	100	one	100
two three four five six	100 100 100 100 100	more than one	500

Table A Table B

In Table A we have listed six possible outcomes and indicated that each of these outcomes would be expected to occur **about** _______ times in 600 rolls of the die. 100

20. Statisticians use the most frequently occurring value in a distribution to characterize the central tendency of that distribution. Statisticians call this most frequent value the **mode** of the distribution. Thus, the value having the highest frequency in a frequency table ________ (would/would not) be called the mode of the distribution represented in that frequency table. — would

21.

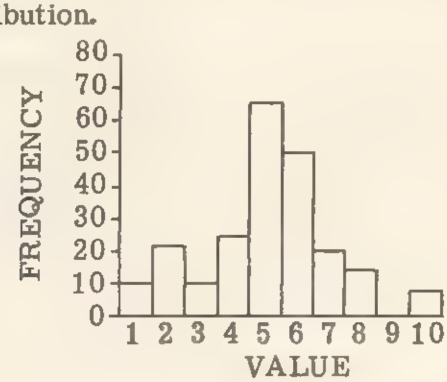
Subject	Score
1	40
2	20
3	10
4	40
5	40

The value occurring most frequently in this collection of data is ______, since this value has a frequency of ____. — 40, 3

22. The **mode** is the most frequently occurring value in a distribution; therefore, 40 is the ________ of the distribution shown in the previous table. — mode

23. The number ______ could be used to characterize the central tendency of the distribution shown in the previous table if you used the mode to characterize the central tendency. — 40

24. The ____________ column on a frequency distribution graph indicates the most frequently occurring value in that distribution. — highest (tallest)



120. (Continued)

of sample space we have been considering and is nothing more than a convenient mathematical way of summarizing a characteristic of the random sampling process. For example, let's consider how this **addition rule** could be applied to rolls of a die. We have seen how certain properties of successive rolls of a fair die can be characterized by a probability distribution. Assigning each of the six outcomes the probability $\frac{1}{6}$ indicates that we would expect each of these six outcomes to occur about __________ of the times we rolled the die in a very large number of rolls. $\frac{1}{6}$

121. The preceding probability distribution indicates that each of the six possible outcomes in the s_________ s________ has been assigned the probability ________ to indicate the limiting relative frequency (or proportion) of times each of these outcomes would be expected to occur in an unlimited number of successive rolls of the die. sample space $\frac{1}{6}$

122. You could represent the probability of obtaining a single dot (a one) on any roll of the die by the symbol **Pr(one dot)**. Similarly, you would represent the probability of obtaining two dots by the symbol ______________. Pr(two dots)

123. Let's consider how the simple **addition rule** would be used to define the probability for a new sample space consisting of the following two outcomes:

Outcome 1:

a single dot

Outcome 2:

more than a single dot

In the new sample, in other words, we have grouped five of the outcomes in the previous sample space containing

24. (Continued)

The value having the highest column in the graph shown above is _____. We could characterize the central tendency of this distribution by reporting that its _________ was 5. — 5 / mode

25. You can indicate that 5 is the most frequently occurring value in the previous data by saying that 5 is the **modal** value. Thus, if the mode of a distribution is 12, you would say 12 was the m__________ (most frequently occurring) value in that distribution. — modal

26. Suppose you were considering a distribution in which the largest frequency was 10. Suppose, however, that more than one value had a frequency of 10. In this case, **both** of these values could be referred to as the mode of that distribution. For example, suppose you had a collection of data consisting of 10 observations of a variable with 3 possible values: a, b, and c. If the frequency of both a and b was 4, and the frequency of c was 2, you ____________ (could/could not) say that **both** a and b were modes in that distribution. — could

27. Consider the table of data shown below.

Observations	Value
1	11
2	11
3	21
4	11
5	21
6	33
7	11
8	21
9	21
10	33

The value 11 has a frequency of ______. — 4

The value 21 has a frequency of ______. — 4

The value 33 has a frequency of ______. — 2

119. (Continued)

group the opinions of all those students who were in favor of the overseas campus. We decided not to distinguish between opinions that represented the same point of view concerning the overseas campus. In the sample space consisting of 10, 000 outcomes, we assigned the probability 1/10, 000 to each outcome. In the new sample space consisting of the **grouped** outcomes, we assigned the probability 1/10 to the outcome "unfavorable opinion" and 9/10 to the outcome "favorable opinion." The probability 1/10 was appropriate since 1, 000 of the 10, 000 students had an unfavorable opinion and, therefore, if each particular student's opinion tended to occur equally often, the unfavorable opinion would tend to occur about ________ 1, 000
times out of every 10, 000 opinions sampled.

Whenever we ask what is the probability of obtaining one of a particular **group** of outcomes in a sample space, we are really asking the question what is the probability of obtaining the opinion of one member of the group, or the opinion of another member of the group, or of another member of the group, and so on.

120. The **rule** of probability theory that can be applied when we assign a probability to a group of outcomes is called the **simple addition rule**. According to this rule, the probability of obtaining any one outcome in the group of outcomes is simply the total of the probabilities assigned to the individual outcomes in the group. This simple **addition rule** of Probability Theory applies to the type

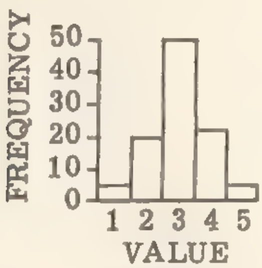
27. (Continued)

The two modes in this collection of data are the values ______ and ______. 11, 21

28 It would be possible to find the value that occurred more frequently than any other value in either numerical or non-numerical data.

Therefore, a collection of non-numerical data can be described in terms of the modal, or most frequently occurring value. For example, if you tossed a coin 5 times and observed 3 "heads" and 2 "tails," you could say ______________ (heads/tails) was the **modal** (most frequently occurring) value in your data. heads

29. Although the mode is often a useful way of characterizing the **central tendency** of a distribution, it is sometimes misleading to describe a distribution by its mode. Consider the two frequency distribution graphs shown below:



GRAPH A

GRAPH B

Notice how the mode of Distribution _____ (A/B) is near the center of the distribution and is typical of the values in that distribution, whereas the mode of the other distribution is not near the center of the distribution and is not particularly typical of the values observed in that collection of data. A

116. This suggests the probability distribution appropriate for our new sample space (the space consisting of the two outcomes "favorable opinion" and "unfavorable opinion") would be one in which the probability 9/10 was assigned to the outcome "favorable opinion" and the probability ________ assigned to the outcome "unfavorable opinion." 1/10

117. Assigning the probability 1/10 to the outcome "unfavorable opinion" simply represents a characteristic of the random sampling procedure for obtaining a single opinion from this particular population. It is consistent with the **limiting relative frequency** procedure we discussed earlier, since if we were to repeat the sampling procedure over and over again, the proportion of unfavorable opinions obtained in this manner would tend to approach the value ____________. 1/10

118. It is **as if** the list of outcomes for repeated draws of a single opinion from the jar containing 10, 000 opinions were viewed as a sample from an **unlimited** population consisting of an unlimited number of draws of a single opinion. As this sample (list of outcomes for successive draws) becomes larger and larger, the proportion of unfavorable opinions in this sample tends to approach a steady value of 1/10. This is all that is meant by assigning the probability __________ to the outcome 1/10
"unfavorable opinion" and the probability ________ to 9/10
the outcome "favorable opinion".

119. Whenever you **group** together the outcomes in one sample space to form a new sample space (as we did in the previous illustration), you can apply a simple rule of probability to assign probabilities to the new sample space. In the previous illustration, we decided to

30. Partly because the mode of a distribution is not always representative of the central or typical value in a distribution, statisticians have defined other ways of characterizing the central tendency of a distribution.

Another way of representing central tendency is to report a value which is **smaller** than the same number of observations as it is larger than. This value is called the **median** of the distribution. For example, suppose your data consisted of the observations 4, 5, 7, 8, and 10. The value 7 would be greater than ____ of the remaining observations and smaller than ____ of the remaining observations. Therefore, we could call ____ the **median** of these five observations.

3/2 | 2
3/2 | 2
7

31. The easiest way of finding the median of a collection of data is to list all the observed values in the order of their size. This procedure is called **ranking** the data. If your data consisted of the values 4, 3, 8, and 7, then the list of values 8, 7, ____, and ____ would be a ranking of the data.

3/4 3/4 | 4, 3

32. If you were to rank the values 10, 6, 11, and 4, you would start with the largest value and end with the smallest to form the list ____, ____, ____, and ____.

11, 10, 6
4

113. (Continued)

simply decided **not** to distinguish between particular students' opinions if their opinions are the same. In other words, we have **grouped** the opinions of all those students **in favor** of the overseas campus as one outcome, and the opinions of all those students who ____________ (are/are not) in favor of the overseas campus as the other outcome. are not

114. For purposes of illustration, let's suppose only 1,000 of the 10,000 students were **opposed** to the overseas campus. We have already indicated that if the sampling procedure were truly random each **particular** student's opinion would be chosen about the same number of times as the sampling procedure was repeated over and over again. That is why we assigned the probability 1/10,000th to each **particular** student's opinion. However, while each particular student's opinion would be sampled equally often, you would expect to obtain **more** opinions ____________ (in favor of/opposed to) the overseas campus by repeated sampling, since only 1,000 of the 10,000 students were opposed to the overseas campus in the population. in favor of

115. While each particular student's opinion would tend to be sampled about the same number of times, only about 1,000 out of every 10,000 opinions sampled (one out of every ten) would be opposed to the overseas campus, since only 1,000 of the 10,000 students' opinions were unfavorable. In other words, if you repeated the sampling procedure over and over again, you would find about 1,000/10,000 or 1/10 of the opinions sampled were unfavorable and about ______/10 of the opinions were favorable. 9

33. You could rearrange the observations in a table of raw data, arranging the values in the order of their size rather than in the order in which they were observed. Thus, we would be forming this new table by ranking the observations.

Notice that the data shown in Table B ____________ (were/were not) formed by ranking the data in Table A. were not

Observation	Value
1	10
2	12
3	11
4	8

TABLE A

Rank	Value
1	12
2	11
3	8
4	10

TABLE B

34. Table D (shown below) ____________ (was/was not) formed by ranking the data in Table C. was

Observation	Value
1	10
2	12
3	11
4	8

TABLE C

Rank	Value
1	12
2	11
3	10
4	8

TABLE D

35. We stated earlier that the **median** value in a collection of data was a value which was ____________ than half of the other observed values and ____________ than the other remaining values. smaller larger

36. One way of finding the median of a collection of data is to **rank** the data and locate the value which divides this list of ranked values in half. If 10, 7, 6, 2, and 1 were a list of ranked values, the value _____ would divide the list in half. **Two** of the other values would be larger than 6, whereas __________ (number) of the remaining of the other values would be smaller than 6. 6 2

111. To express this in terms of relative frequency, you would expect the proportion of times each particular student opinion was sampled (as you repeated this procedure over and over again) to be about the _______. same

112. To summarize, if the procedure by which you obtained a single student opinion from the population of 10, 000 student opinions were truly random, you would expect the proportion of times each particular opinion was obtained in repeated applications of this random sampling procedure to be about the same. Since there are 10, 000 opinions, the proportion of times each student's opinion could be expected to be sampled in a large number of repeat samples of one opinion each would be about _______________________. Each opinion would tend 1/10, 000th

1/10, 000th - 2/10, 000ths

to be sampled about 1/10, 000th of the time in repeated samples of a single opinion, since the total of these proportions would equal 1 and since each student's opinion would be sampled the same proportion of times. This suggests that according to the **limiting relative frequency** procedure for assigning probabilities to a sample space, the number (probability) ____________ should be assigned to each of the 10, 000 possible outcomes in the sample space. 1/10, 000th

113. We could define another sample space for the process of randomly drawing a single opinion from the population of 10, 000 opinions. Suppose you defined a sample space consisting of the two outcomes: a favorable opinion and an unfavorable opinion. This would be a sample space, since each opinion sampled would correspond to one and only one of these two outcomes. Actually, we have

36. (Continued)

Since 6 is larger than half of the remaining values and smaller than the other half, the **median** would be _____. 6

37. It is a simple matter to find the median of a distribution when you have an odd number of observations. You simply rank the observations and find the middle value in this list of ranked observations. This middle value would be the ___________ of your data. median

38. If you had an even number of observations, there would not be a value in the list of ranked data such that the same number of observations fell above and below that value.

To illustrate this problem consider the table of ranked data shown below:

Rank	Value
1	80
2	60
3	40
4	30
5	20
6	10

Notice how we have indicated the rank of each value in the first column of the table. The largest observed value has a rank of ____ (1/6) and the smallest observed 1

value has a rank of ____ (1/6). 6

Notice also that there are ____ observations larger 2

than the value 40 and ____ observations smaller than the 3

value 40. Thus, 40 ________ (is/is not) the median. There are __ is not, 3

observations larger than the value 30 and ___ observations 2

less than the value 30, indicating that 30________ (is/is not) the median. is not

107. We have now considered how probability distributions can characterize observations from an infinite population. Let's consider how a probability distribution could be defined characterizing the process of drawing samples randomly from a finite population. Earlier, we considered an example in which the opinions of 10,000 college students, each of which was either for or against an overseas campus, were all written on slips of paper. These slips of paper were placed in a basket, thoroughly mixed, and then withdrawn to form a sample of student opinion. This process was represented as a r________ procedure for obtaining a sample of student opinion from the ____________ population of 10,000 students' opinions.
(finite/infinite)

random

finite

108. The population of student opinions would be **finite** since there was a ________________ number of opinions in the population.
(limited/unlimited)

limited

109. Suppose you defined a **sample space** of all possible outcomes for the random process of drawing a single student opinion from the jar of 10,000 opinions. There would be 10,000 possible outcomes, one outcome for each **particular** student opinion. Therefore, a sample space could consist of _______ outcomes, one for each particular student opinion.

10,000

110. According to our definition of a random sampling process, each of the 10,000 student opinions would have the same opportunity of being included in the sample. Therefore, if the random process by which we obtained a single student opinion was repeated over and over again for a very great number of times, you would expect each particular student's opinion to be sampled about the ______________ number of times as would any other student's opinion.

same

38. (Continued)

Neither 30 nor 40 is the median, since too many values are smaller than ______ (30/40), whereas too many values are 40

larger than ______ (30/40). 30

39. Strictly speaking, any value between 30 and 40 could be called the median of this data. However, statisticians have agreed upon a rule for finding the median of an **even** number of observations. They would say that the median of the previous collection of data was a value halfway between 30 and 40. In other words, 35

______________ (would/would not) be called the median, because would

35 is halfway between 30 and 40.

40. Suppose 8, 6, 4, and 3 were ranked data. The value halfway between 6 and 4 would be the value ______. 5

Therefore, 5 would be the ____________ of this data. median

41. Three lists of ranked data are shown below.

Data A: 5, 4, 1

Data B: 6, 3, 2

Data C: 5, 3, 1, 1

The middle value in Data A is the value 4; therefore, 4

would be the ______________ of Data A. median

The middle value in Data B is the value _____. This value, 3

therefore, would be the ________ of that data. median

104. The proportion of times a baseball player has gotten a hit out of all the times he has officially been to bat is often referred to as his batting average. In other words, if he has been to bat 10 times and gotten 3 hits, his batting average would be .3 or, as it is often written, .300. If a batter had a batting average of .400, this would mean that he had gotten a ________ on exactly ______________ of his times at bat. — hit / 4 tenths

4-tenths/400

105. Suppose you viewed the process that determined whether or not a particular batter obtained a hit as a **random process**. You might view each time at bat as an additional observation from an infinite population consisting of an unlimited number of times at bat. Suppose the batter had been at bat a great number of times and had obtained a hit about 3-tenths of the time. You might represent his performance as a probability distribution. The **sample space** to which you could assign probabilities would consist of the two mutually exclusive and exhaustive outcomes of a time at bat: a hit or ______ hit. (We will consider anything other than a hit as "no hit".) Since the proportion of hits by this particular batter approached the value .3 after a great number of times at bat, you could assign the probability .3 to the outcome "hit" and the probability ____________ to the outcome "no hit." — no / .7

106. The probability of a hit (.3) is a way of characterizing the batter's performance if you view his performance as a r__________________ process. This probability implies that the proportion of hits in a very large number of observations (official times at bat) appears to approach the value _______. — random / .3

41. (Continued)

Notice that Data A and B consist of an ______ number (odd/even) of observations, whereas Data C consists of an ______ number of observations. Using the rule for (odd/even) data consisting of an even number of observations, we would find the value halfway between ___ and ___; this is the value _____. Therefore, we would say the median of Data C is ____.

odd

even

3, 1

2

2

42. Consider the table of data shown below:

Observation	Value
1	100
2	200
3	160
4	30
5	180

The largest value in this table is ______ and the smallest value is _____.

200

30

43. If we **rank** the data in the preceding table, our list of values ____________ (would/would not) be 200, 160, 180, 100, 30.

would not

44. The previous list was incorrectly ranked. The value "180" should have come before ______ rather than after it, since 180 is larger than _____.

160

160

45. We refer to the largest value in a table of ranked data as having rank 1. The value having rank 1 in the previous collection of data would be _____. The value having rank 2 would be _____.

200

180

101. You have just seen that the relative frequency or ________ of observations of a particular value in a random sample tends to vary less and less as the size of the sample is increased. — proportion

102. This **limiting relative frequency** can be thought of as the population proportion of an infinite population. Since an infinite population consists of an unlimited number of observations, we could never calculate the actual population statistic. It is often useful, however, to view the sample statistic in a very large sample **as if** it were the population statistic. This is reasonable, because we know the **difference** between a sample statistic and the population statistic in a very large sample will tend to be quite ________ (large/small) according to the law of of l________ n________. — small; large numbers

103. A **probability distribution** can represent an infinite population consisting of an unlimited number of outcomes produced by a random process. You have seen, for example, how a p________ distribution can represent an infinite population consisting of an unlimited number of tosses of a coin. When you say that the probability of heads equals one-half, you simply mean that the limiting r________ f________ of heads is one-half as the sample size is made very large. — probability; relative frequency

46. The **median** of the previous collection of data is the value ______, since the values ______ and ______ are larger and the values ______ and ______ are smaller than this median value. — 160, 180, 200; 100, 30

47. However, if the values in the previous collection of data were 200, 180, 160, and 100, the median would be a value halfway between _____ and ______. — 160, 180

48. To find the difference between 180 and 160, you subtract 160 from 180. Half of this difference is ______, which is what you would add to the value 160. Therefore, the median is ______. — 10; 170

49. We have now considered **two** ways of representing the central tendency of a distribution: by the _________ and by the ____________. — median; mode

50. The ____________ (mode/median) of a distribution is the most frequently occurring value in the distribution, whereas the ____________ (median/mode) is a value that is less than half of the other values and greater than half of the remaining values. — mode; median

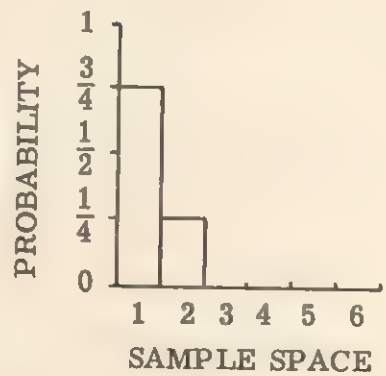
51. Suppose your data were the values:

10, 5, 1, 10 and 4.

The median of the distribution is the value ______, whereas the mode is the value ______. — 5; 10

99. It is important to keep in mind that probabilities assigned in this fashion characterize the **process** rather than any particular outcome of the process. In other words, the probability $\frac{1}{2}$ assigned to the outcome "heads" for repeated tosses of a **fair** coin does not indicate what will happen on any particular toss. What it represents is the tendency for the proportion of heads in a large number of tosses to approach the value _______. This is a characteristic **not** of one toss but of the whole series of tosses — that is, the **process** of tossing a fair coin. $\frac{1}{2}$

100. This interpretation of probabilities is often referred to as the **limiting relative frequency** concept of probability. As a simple illustration of another procedure for assigning probabilities, consider the following example. Suppose you were about to roll a die and you asked someone to assign a number to each of the six possible outcomes, such that the sum of numbers totaled 1 and the size of each number was related to his "**subjective degree of certainty**" that a particular outcome would occur. In other words, if he had a hunch that 6 dots was more likely to occur than 2 dots, he should assign a larger number to the outcome 6 than to the outcome 2. The probability distribution shown below could be used to characterize a possible assignment of numbers by this procedure.



According to this assignment, the subject expected the number of dots showing on the die to be ____ or ____. 1, 2

52.

Observation	Value
1	10
2	9
3	10
4	11
5	7
6	10
7	5
8	2
9	1
10	5

The mode of the collection of data shown above is _____ since this value occurs _____ times.

10

3

53. The value 9 ________ the median of the previous
is/ is not

collection of data since there are _____ observations with values larger than 9 and _____ observations with values smaller than 9.

is not

4

5

While 9 is too large to be the median, 7 is too _________ to be the median.

small

54. To find the median of the previous data, you would find the value halfway between ______ and ______. Therefore the median of the previous collection of data is _____.

9, 7

8

55. The median (like the mode) can sometimes give a misleading picture of a distribution. For example, consider the two collections of data shown below:

Data A: 100, 99, 98, 97, 96.

Data B: 100, 99, 98, 4, 2.

The median of Data A is ____ and the median of Data B is ______.

98

98

95. According to **Graph A**, Pr (one dot) = Pr (two dots) = Pr (three dots) = Pr (four dots) = Pr (five dots) = Pr (six dots) = ______. $\frac{1}{6}$

96. Thus, the **probability distribution** in Graph A represents a characteristic of repeated rolls of a die. The tendency for the proportion of each possible outcome to approach a value of $\frac{1}{6}$ as the number of observations becomes larger is the feature of this particular random process, characterized by the ______________ ______________ shown in Graph A. probability distribution

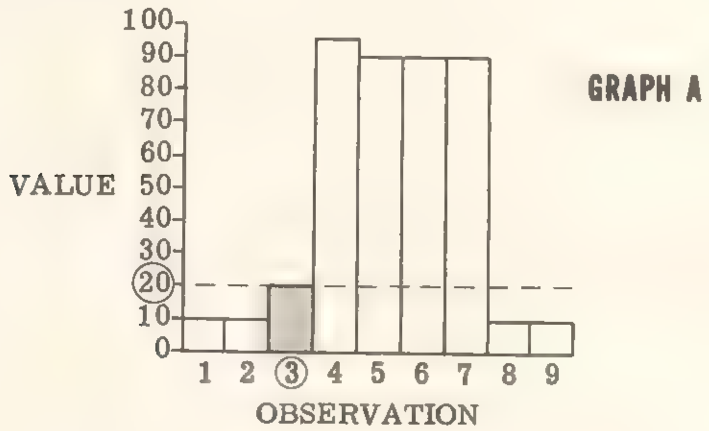
97. The probability distribution shown in the Graph B (above) would suggest that, as the number of rolls of the die was increased, the proportion of times one dot was rolled tended to approach the value ___, while the proportion of times that two dots were rolled tended to approach a value ______________ (greater than/less than) $\frac{2}{6}$. $\frac{2}{6}$ less than

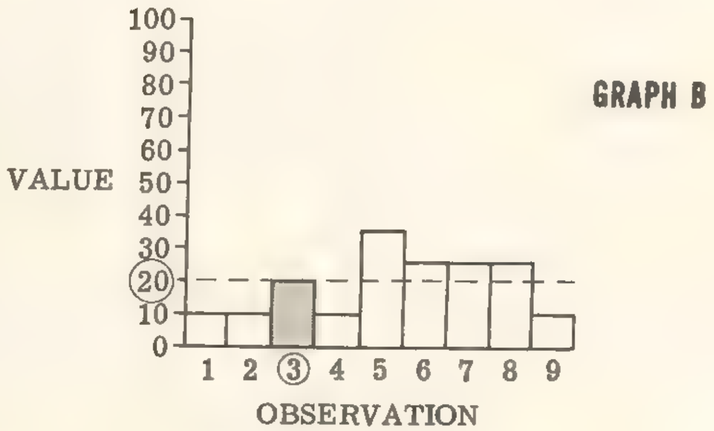
98. The die that behaved in this manner would not be consistent with the usual idea of a **fair** die. If a die were made properly and rolled properly, you would generally expect that each of the six possible outcomes would occur about the same proportion of times in the long run. Thus, the probability distribution in Graph ______ (A/B) could be thought of as describing the behavior of a **fair** die, whereas the probability distribution shown in Graph ______ (A/B) suggests the behavior of a **biased** die, one which tended to produce a one-dot or six-dot roll more often than the other possible outcomes. A B

56. While both distributions have the same median, the values below the median in Data ______ differ much more from the median than do the values above it. B

A/B

57. The median only indicates the value dividing the list of ranked data into two equal parts. The median does not indicate **how much** smaller or **how much** larger are the values falling above it or below it in the list. This can be illustrated by the following raw data graphs.





Notice that in **both** collections of data ______ observations had a value **larger than** the value of observation 3 and ______ observations had a value **smaller than** the value of observation 3. Thus, ______ is the median of both distributions since it is the value of observation ______.

4

4

20

3

90. (Continued)

The Graph A indicates that the first roll of the die resulted in one dot showing. Thus, on this graph "one dot" has a proportion of 1 and all the other outcomes have a proportion of _______ after the first roll of the die. 0

91. After 6 rolls (Graph B) the only outcomes that had not occurred at least once were 2, 4, and _____ dots showing. 6

92. After 60 rolls (as indicated on Graph C) the proportions of times each of the 6 possible outcomes had occurred were more similar. After 600 (Graph D) rolls each of these proportions was very close to _________. $\frac{1}{6}$

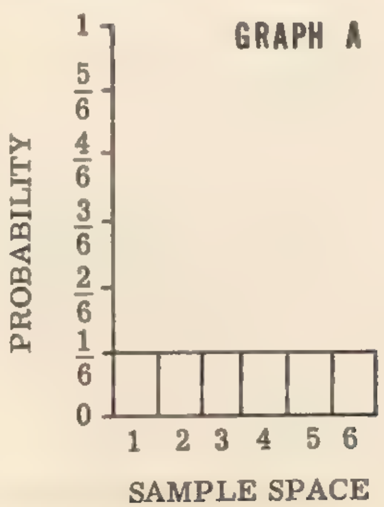
$\frac{1}{6}$ / $\frac{2}{6}$

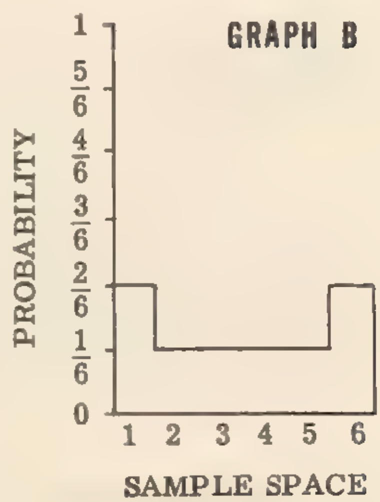
93. Finally, after 6, 000 rolls (Graph E) all 6 sample proportions were extremely close to _______. $\frac{1}{6}$

94. You could represent this characteristic of repeated rolls of a die by the probability distribution shown in Graph _____. It indicates that the A

A/B

proportion of each of the possible outcomes tended to approach a value of $\frac{1}{6}$ as the number of observations was increased.





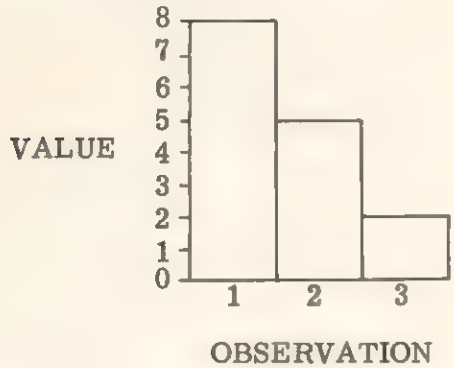
58. The table of data shown below contains ______ observations of a ______________ variable. 3 numerical

numerical/non-numerical

Observation	Value
1	5
2	8
3	2

59. The graph of raw data shown below ______________ represent the data in the previous table. does not

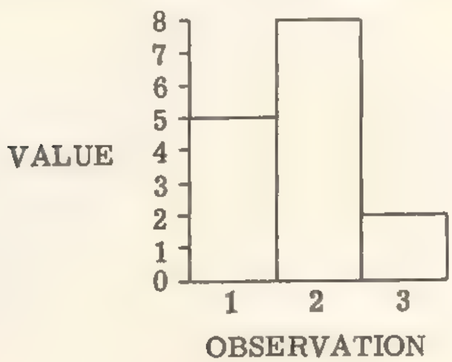
does/does not



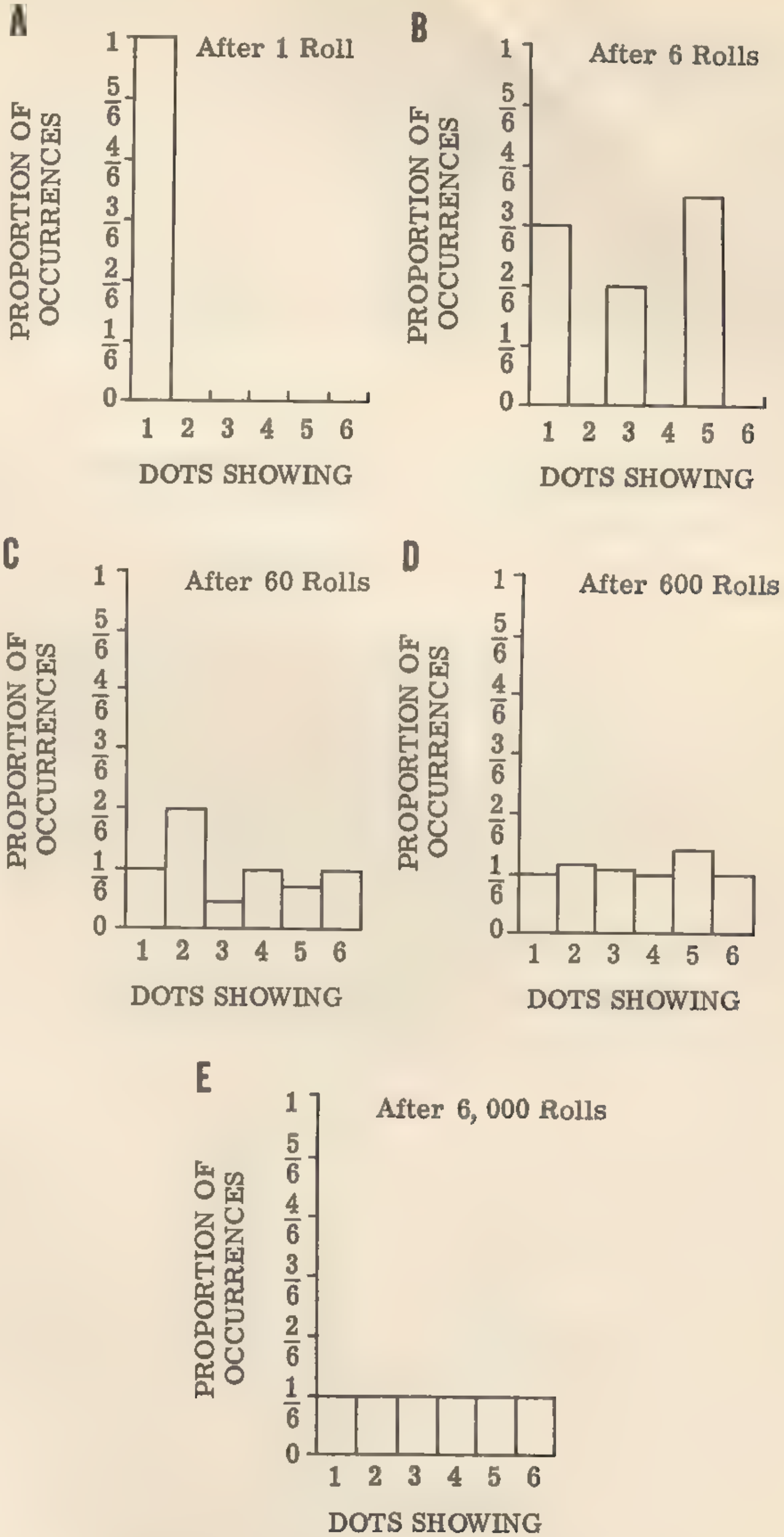
60. The previous graph does not represent the previous table of data, since observation 1 had a value of 5 and observation 2 had a value of 8 in the table, whereas observation 1 had a value of ______ and observation 2 had a value of _____ in the graph. 8 5

61. The graph of raw data shown below ______________ represent the data shown in the previous table. does

does/does not



90. (Continued)



61. (Continued)

All the observed values in the previous graph ________ (are/are not) the same. — are not

Differences in the observed values are indicated on the graph by differences in the ____________ of the 3 columns. — heights

Thus, the value of observation 1 is clearly less than that of observation 2, since the height of column 1 is less than that of column 2. Similarly, the value of observation 2 is ___________ (less/greater) than the value of observation 3, since column ______ is higher than column ____. — greater than; 2; 3

62. While you can always compare the value of one observation to the value of another observation, it wouldn't make much sense to simply say, "Observation 3 is less than." The immediate question would be: "Less than what?" It is often convenient to pick a **reference value** for comparison with the observed values. For example, suppose you chose the value 4 as a reference value. You could then describe the data in the previous graph by saying, "Observation 1 is greater than 4, observation 2 is greater than 4, and observation 3 is _________ than 4." — less

87. (Continued)

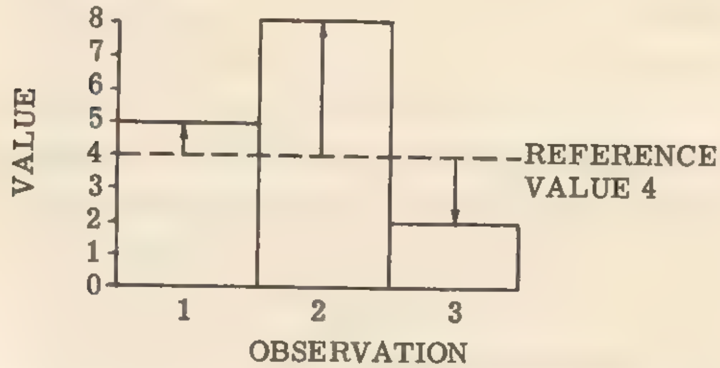
a value of .5. In other words, it appears as if the successive tosses of the coin were all randomly drawn from an infinitely large population in which the population proportion of heads equaled _____. .5

88. Notice that if the proportion of heads tends to approach a value of .5 as the number of observations is increased, the proportion of tails must approach a proportion of _______ (since the total of the two proportions must .5
equal ______). 1

89. A **probability distribution** can be defined which represents this characteristic of repeated tosses of a coin. You simply assign the number .5 to the outcome "heads" and the number ______ to the outcome "tails." .5

90. This probability distribution **characterizes** something interesting about repeated tosses of a coin — namely, the tendency for the proportion of heads and tails to approach the values indicated in the probability distribution as the number of tosses (observations) of the coin is increased. Let's consider a similar illustration, only this time, instead of successively tossing a coin, imagine successively rolling a die. Suppose the die were rolled 6,000 times. You could calculate a proportional distribution of the 6 possible outcomes (one dot, two dots, etc.) after one roll, after six rolls, 60 rolls, 600 rolls, and after 6,000 rolls. The following proportional distribution graphs indicate the manner in which these sample proportions tend to approach a steady value as the number of observations is increased.

63. The relationship between the reference value 4 and the value of each of the three observations is made clear by the following graph.



While this graph represents the same data as does the previous graph, we have indicated the reference value 4 with a dotted line drawn across the graph at a height equal to the value _____. We have also indicated the difference between the reference value and each observed value with an arrow. The arrow in column 1 is pointing upwards, since the observed value 5 is greater than the reference value 4. Similarly, the arrow points ________ (up/down) in column 2, since the value of observation 2 is ___________ (greater/less) than the reference value 4.

4

up

greater

64. The value of observation 3 is ___________ (less/greater) than the reference value 4. Therefore, the arrow on our graph points ________ (up/down). Notice that the **difference** between the value of observation 1 and the reference value is less than the difference between the value of observation 2 and the reference value. This is indicated on the graph by the fact that the arrow in column 1 is ____________ (longer/shorter) than the arrow in column 2. Thus, the **size of the difference** between the observed value and the reference value is indicated by the ____________ of the arrow representing that difference.

less

down

shorter

length

85. Looking down the column of sample proportions (column 3), it is clear that as more and more observations were included in the sample, the sample proportion tended to change less and less and seemed to settle near a value of ______. .5

.4/ .5

86. The fact that a sample statistic (such as a proportion) tends to approach a steady value as the size of the sample is increased is an example of the law of l__________ large

n________________. It is **as if** the sample statistic were approaching closer and closer to the value of the population statistic as you increased the size of the sample. numbers

It is this tendency of the sample statistics of random samples from an infinite population to approach a steady value as the sample size is increased that allows us to define the population proportion. The population proportion of an infinite population is defined as that value which the sample statistic tends to approach as the size of the sample is increased. According to the previous table, therefore, it would appear that the population proportion in the population consisting of an unlimited number of tosses of a coin was approximately

________. .5

.5 / .4

87. This characteristic of repeated tosses of a coin allows us to say something about the uncertainty concerning the outcome of any particular toss. While you cannot predict exactly what will occur on any particular toss, you can say that as the number of tosses is increased, the proportion of heads seems to grow closer and closer to

65. The length of the arrow represents the difference between the reference value and the observed value, regardless in which direction the arrow is pointing. Thus, the arrow points up in column 2 since the second observed value is greater than the reference value, whereas the arrow points down in column 3 because the third observed value is __________ than the reference value. However, the value of observation 3 is closer to the reference value than is the value of observation 2. This is indicated by the fact that the arrow in column _____ is ____________ than the arrow in column 2.
shorter/longer

less

3

shorter

66. Differences between observed values and a reference value (represented with arrows in the previous graph) are referred to as **deviations** by statisticians. If a value is greater than the reference value, that value has a **positive deviation** from the reference value. If a value is less than the reference value, that value has a **negative deviation** from the reference value. Since the value of observation 1 (on the previous graph) is greater than the reference value 4, you would say that the value of observation 1 had a positive deviation from the reference value 4. Similarly, since the value of observation 2 is ____________ than the reference value 4
greater/less
you would say the value of observation 2 had a ______________ deviation from the reference value 4.
positive/negative

greater

positive

Since the value of observation 3 is less than the reference value 4, we would say the value of observation 3 had a _______________ deviation from the
positive/negative
reference value 4.

negative

81. (Continued)

This table indicates what you might observe if a coin were tossed repeatedly 10, 000 times. If on the first toss a **tail** occurred, the outcome could be viewed as a sample of size 1 from an infinite population consisting of an unlimited number of tosses.

The sample proportion of **heads** would be 0, since no **heads** had occurred in 1 observation. This fact is indicated in the first row of the preceding table. The number 1 in the first column indicates the number of **tosses**. The 0 in the second column indicates the number of **heads** that have occurred in that many tosses, and the 0 in the third column indicates the sample proportion at that point. The second number in the first column is a
2, which indicates that ______ tosses have been made. 2

82. Note that in the first two tosses the number of heads in
the sample was ______, which represented a sample 0
proportion of heads equal to ______. 0

83. After five tosses, however, the number of heads observed
was ______, which represented a sample proportion of 3
______. After ten tosses three heads had been observed, .60 ($\frac{3}{5}$)
which led to a sample proportion of ______. .3 ($\frac{3}{10}$)

84. Finally, when all 10, 000 tosses had been made, as
indicated in the last row of the table, exactly ________ 5, 028
heads had occurred, which represents a sample
proportion (rounded off to the nearest 100^{th}) of ______. .50

67. The amount by which a particular value deviates from the reference value is indicated by the **length** of the arrow in the previous graph. Whether the deviation is positive or negative is indicated by the **direction** in which the arrow points. In other words, whenever an observed value is greater than the reference value, the arrow will point ________ (up/down), indicating a ________________ (positive/negative) deviation. Whenever the observed value is less than the reference value, the arrow will point ________ (up/down), indicating a ________________ (positive/negative) deviation.

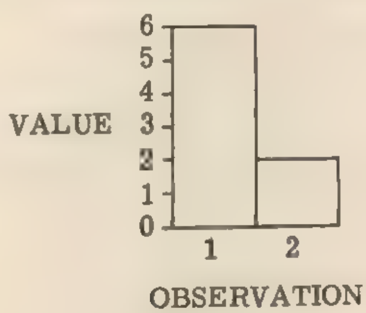
up, positive

down

negative

68. The graph of raw data shown below represents 2 observations having a value of _____ and _____.

6, 2



69. Suppose you chose 0 as the reference value for describing the two observations in the previous graph. Both of the observed values would be ____________ (greater/less) than the reference value and would represent ________________ (positive/negative) deviations from the reference value.

greater

positive

79. (Continued)

outcomes for 10 tosses of a coin as a sample from this infinite population. You could also view the list of outcomes for 20 tosses as a larger __________ from this infinite ______________.

sample

population

80. According to the law of large numbers, you would expect the sample proportion, p, to be closer to the population proportion, $\tilde{p}$, for a __________ (large/small) sample rather than for a __________ (large/small) sample.

large

small

81. As you increase the size of a random sample from an infinite population, a statistic (such as a sample proportion) will appear to grow closer and closer to some particular value. This can be illustrated by the following table.

Figure 6

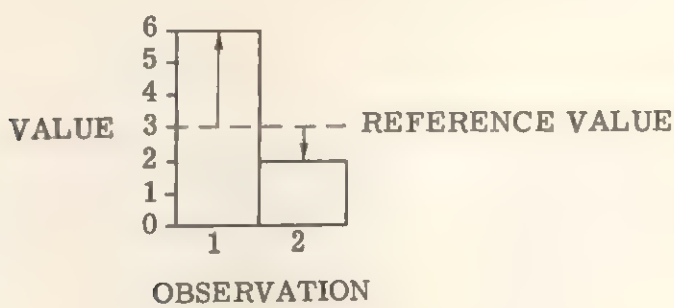
Sample proportion tends to approach a steady value as sample size is increased.

Number of tosses completed	Number of heads observed	Proportion of heads in sample
1	0	0
2	0	0
5	3	.60
10	3	.30
50	20	.40
100	48	.48
500	260	.52
1, 000	510	.51
5, 000	2, 542	.50
10, 000	5, 028	.50

70. Suppose you choose 8 as a reference value. Since the value of both observations is ___________ than 8, both less
greater/ less
deviations would be _______________. negative
negative/ positive

71. Suppose you chose a reference value somewhere between 6 and 2. The value of observation 1 would represent a
_______________ deviation from the reference value, positive
positive/ negative
whereas the value of observation 2 would represent a
_______________ deviation. negative
positive/ negative

72. We have redrawn the preceding graph and indicated below the deviation of each observation from a **reference value** of _____. 3



The reference value is closer to the value of observation 2 than it is to the value of observation 1. This is indicated by the fact that the arrow pointing up is
_____________ than the arrow pointing down. We can longer
longer/ shorter
calculate the actual size of the deviation simply by **subtracting the reference value from the observed value.** For example, you would calculate the deviation of observation 1 from the reference value by subtracting 3 from ______. 6

75. (Continued)

actually lands on one side or the other.) These two outcomes represent a sample space, since the outcomes are m____________ e____________ (only one outcome can occur each time the coin is tossed). The two outcomes are also exhaustive because they represent ________ possible outcomes. — mutually exclusive; all

76. Suppose we tossed the coin repeatedly, each time listing the outcome of the toss. This list would constitute a collection of data because it is a list of observed values of a variable called "outcome," whose two possible values are ________ and ________. — heads, tails

77. Imagine repeating the tossing procedure over and over an unlimited number of times. This would produce an unlimited collection of data. If this collection of data is regarded as a population, it would be described as a(n) ________ (finite/ infinite) population, since it consists of an **unlimited** number of observations. — infinite

78. To determine the proportion of heads for a finite (in other words, limited list of outcomes), you could simply divide the number of ________ observed by the total number of observations. — heads

79. Since you could not possibly collect all of the data in an infinite population, a problem arises concerning how you would define the population proportion, $\tilde{p}$. The law of large numbers provides us with a way of defining this population proportion. You can view a collection of

73. The deviation of observation 1 from the reference value 3 equals 3 subtracted from 6, which equals ____. 3

74. Similarly, you would calculate the deviation of observation 2 from the reference value 3 by subtracting 3 from ____. 2

75. Notice that 2 minus 3 equals -1. This means the deviation of observation 2 from the reference value 3 is ______ (-1/+1). -1

When we subtract a reference value from a smaller value, our answer will be negative. Therefore, all **negative** deviations will be represented by ______________ (positive/negative) numbers. negative

On the other hand, all ______________ deviations will be represented by positive numbers, since the observed value is, in this case, greater than the reference value. positive

76. In the table shown below, we have summarized the information concerning deviations of the observed values from the reference value 3.

OBSERVATION	VALUE	DEVIATION FROM 3
1	6	3
2	2	-1

77. We have simply added another column to the previous table of raw data and listed the deviations of each of the observed values from the value _____. Thus, the numeral -1 in the last row of the third column represents the deviation of observation 2 from the reference value 3. We obtained the deviation of observation 2 from the reference value 3 by subtracting _____ from _____, to give us an answer of _____. 3 | 3, 2, -1

72. This characteristic of random samples is part of what is meant by **the law of large numbers**. The main point of this law is that a sample statistic, such as a proportion, is more likely to be similiar to the population proportion if the size of the sample is large. In other words, if the population proportion were .8, the sample proportion from a sample of size ________ would tend to be closer to .8 than would the sample proportion in a sample of size ________. — 100, 10

10/100

10/100

73. The idea that the sample proportion will tend to be closer to the population proportion the larger the size of the sample is, therefore, the main point of the law of ____________ ______________. This law does not imply that a sample proportion from a large sample is always more similar to the population proportion than is one from a small sample. It says that it ________ to be closer. In other words, it is **more often** closer than is one from a small sample. — large numbers; tends

74. You also saw that a sample mean is more likely to be close to the population mean if the sample is large than if the sample is small. This is another example of the ____________ of __________ ______________. — law, large numbers

75. Let's consider how this characteristic of random samples serves as a basis for assigning probabilities to the members of a sample space. The two outcomes, heads and tails, represent a sample space for a single toss of a coin. (We shall only consider tosses where the coin

78. The following table contains the same two observations. However, we have left room to indicate deviations from the reference value ______ rather than 3 (as in the previous table). 4

OBSERVATION	VALUE	DEVIATION FROM 4
1	6	
2	2	

79. To find the deviation of observation 1 from the reference value 4, you subtract 4 from ______ (6/2). 6

80. Therefore, the deviation of observation 1 from the reference value 4 is ______. 2

81. The following table ______ (is/is not) correct. is not

OBSERVATION	VALUE	DEVIATION FROM 4
1	6	2
2	2	2

82. The previous table is incorrect because observation 2 was **smaller** than the reference value. Therefore, the deviation of observation 2 from the reference value would have to be represented by ______ rather than 2. -2

83. The following table ______ (is/is not) correct. is

OBSERVATION	VALUE	DEVIATION FROM 4
1	6	2
2	2	-2

67. Each of these probabilities can be no larger than _____ nor smaller than _____, and the total of all of these probabilities must equal _____. — 1, 0, 1

68. Any particular assignment of probabilities to members of a sample space is called a ______________ ______________. — probability distribution

69. Up to this point, we have only described some of the basic ideas involved in **Probability Theory**. Earlier we indicated that Probability Theory was a mathematical system which could be used to represent r__________ p__________, such as games of chance. — random processes

70. Although Probability Theory was originally developed to represent games of chance, it has been found to be useful in representing a wide variety of processes. Therefore, the general definition of a random process is any process that can be represented by __________ __________. — Probability Theory

71. Next, we shall consider procedures for assigning probabilities that make the probability distribution a useful representation of an actual sampling process. Earlier, we indicated that a **large** random sample tends to be more representative of the population than does a small random sample. In general, the larger the size of the random sample, the __________ (more/less) similar will be the sample statistic to the population parameter. — more

84. There is something special about using 4 as the reference value. The positive deviation represents the same difference between the observed value and the reference value as does the negative deviation. In other words, if we represented deviations on a graph as we did previously, the length of the two arrows would be __________ (although one arrow would be — the same
the same/different
pointing up and other down).

85. Because the reference value 4 has the unique property of being as close to the value 6 as it is to the value 2, we say that 4 is the **mean** of the values 6 and 2. Since 6 and 2 are the two observed values in the previous collection of data, we could say that the **mean** of that collection of data is _____. — 4

86. If we added the positive deviation from 4 and the negative deviation from 4, our answer would be ______ because — 0
+2 added to -2 equals _____. — 0

87. Another way, therefore, of describing that unique characteristic of the reference value 4, which makes it the **mean** of the previous collection of data, is to say the **sum of the deviations** from 4 equals _____. — 0

88. Suppose your data consisted of 3 observations instead of only 2 — for example, 8, 3, and 1. If you chose 10 as your reference value, all the deviations would be
__________. If you chose 0 as your reference — negative
positive/negative
value, however, all the deviations would be
__________. — positive
positive/negative

OBSERVATION	VALUE	DEVIATIONS FROM 6
1	8	2
2	3	-3
3	1	-5

61. (Continued)

Suppose the die were rolled and 2 dots were showing. In that case, both outcome _____ and outcome _____ in the previous list would have occurred. — A, D

62. Two dots would represent an "even number of dots" and at the same time it would represent "more than one dot showing." Therefore, two of the outcomes in the previous list of four outcomes would have occurred. This list of four outcomes would **not** be suitable as a sample space since more than one outcome could occur at the same time. The list is **exhaustive** because any outcome would be included in the list. Since more than one outcome can occur at the same time, however, the outcomes are not ______________ ______________. — mutually exclusive

63. A list of outcomes can describe a sample space if the outcomes are both e______________ and m______________ e______________. — ex<u>haustive</u>, m<u>utually</u> e<u>xclusive</u>

64. A simple way of stating that a list of outcomes is both ______________ and ______________ ______________ is to say that one and only one of the outcomes must occur. — exhaustive, mutually exclusive

65. The fact that **one** of the outcomes must occur implies that the list is ______________ (exhaustive/ mutually exclusive) and the fact that **only one** can occur implies that the list is ______________ (exhaustive/ mutually exclusive). — exhaustive; mutually exclusive

66. The numbers assigned to the outcomes in a sample space are called ______________. — probabilities

89. To check the deviation of observation 3 in the previous table, you would subtract the value ______ from the value ______. This would indicate that the deviation in the table was ______________ (correct/incorrect).

6
1
correct

90. We mentioned that a reference value is called a **mean** when the sum of the deviations from that reference value equals 0. For the set of data we considered earlier, for example, the deviations from the mean value 4 were +2 and -2, giving a sum of deviations equal to (+2) + (-2), which equals ____.

0

91. Even when the data consist of more than two observations, we can define the **mean** in the same way. In other words, if we **add all the deviations from a particular reference value** and our answer is 0, that reference value is the **mean** of those observations. Consider the previous table of deviations from the reference value 6. To find the sum of the deviations, we would add +2, -3, and -5, which yields an answer of ______.

-6

92. Since the sum of the deviations of our three observed values from the reference value 6 does not equal 0, the value 6 __________ (is/is not) the **mean** of these three observations.

is not

93. We could record the deviations of each observation from the reference value _____ in the following table:

4

OBSERVATION	VALUE	DEVIATIONS FROM 4
1	8	
2	3	
3	1	

We would find the deviation of observation 1 by subtracting ___ from ____, indicating that the deviation of observation 1 from the reference value was ______.

4, 8
4

58. Notice the probability of outcome _______ on Graph B is the smallest possible value of a probability, since it equals ______. A

0

59. Let's review the two characteristics of a list of outcomes which are necessary for the list to be a sample space. First, the list must include **all** possible outcomes. A list of this sort is said to be **exhaustive**, since it exhausts **all** possible outcomes (it is a **complete** list of **all** possible outcomes). The second necessary characteristic is that **only one** of the outcomes may occur at one time. Such a list of outcomes is said to be **mutually exclusive** (the occurrence of one outcome **excludes** the occurrence of any of the other outcomes since only one outcome may occur at a time). Therefore, a **sample space** consists of a list of outcomes that are both exhaustive and mutually exclusive. The list is ______________________ (exhaustive/ mutually exclusive) if it includes all possible outcomes, and it is ______________________ (exhaustive/ mutually exclusive) if only one of the outcomes can occur at a time.

exhaustive

mutually exclusive

60. A list of outcomes for rolls of a die consisting of the two outcomes "one dot showing" or "two dots showing" could be described as **mutually exclusive**, but it could not be described as e______________, since **all** possible outcomes are not included. exhaustive

61. On the other hand, consider the following list of outcomes for a role of a die.

Outcome A: an even number of dots showing

Outcome B: an odd number of dots showing

Outcome C: one dot showing

Outcome D: more than one dot showing

94. In the same way, we would find that the deviation of observation 2 was ____ and that the deviation of observation 3 was ______. -1 -3

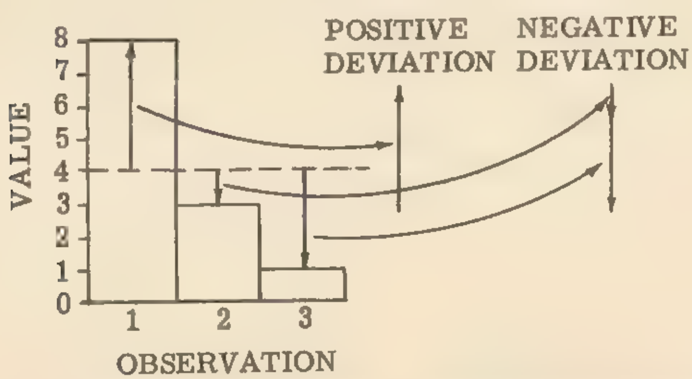
95. We have recorded the deviations from the reference value 4 in the following table:

OBSERVATION	VALUE	DEVIATIONS FROM 4
1	8	+4
2	3	-1
3	1	-3

We said that the **mean** or **average** of a group of numerical observations is that particular reference value yielding deviations whose total is 0. To find the total of the deviations from the reference value 4 for the three observations in this table, we would add ____, ____, and ____. 4, -1 -3

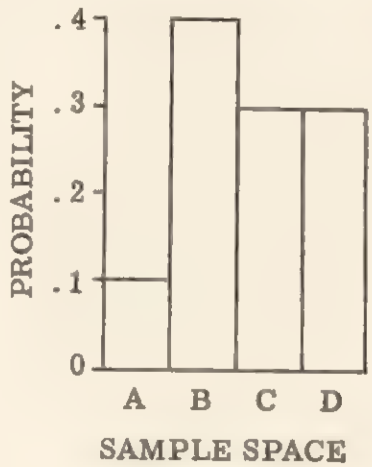
Thus, 4 ________ (is/is not) the mean of this group of observations. is

96. We can illustrate this graphically as follows:

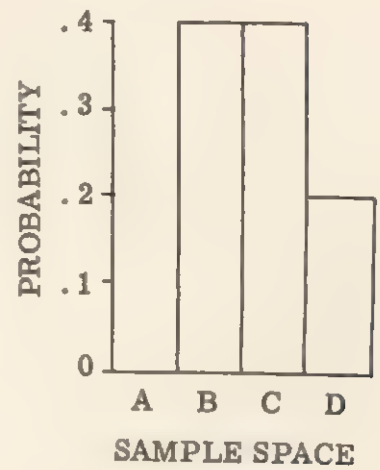


The arrow labeled "positive deviation" represents the one positive deviation in the graph. The two downward pointing arrows, which are connected together, represent the two negative deviations in the graph. In order for the sum of the deviations to equal 0, the positive deviations

55. One of the two graphs shown below could **not** represent a probability distribution. Try to find that graph.



GRAPH A



GRAPH B

You would be correct in saying Graph ______ (A/B) could **not** represent a probability distribution. — A

56. Graph A could not represent a probability distribution since the sum of the four probabilities does not équal 1, as it must in any probability distribution. According to **Graph A**, the probability of A equals .1, the probability of B equals .4, the probability of C equals .3, and the probability of D equals ______. Therefore, the sum of these four probabilities would equal ________. — .3; 1.1 $(1\frac{1}{10})$

57. On the other hand, Graph B **could** represent a probability distribution, since

Pr(A) + Pr(B) + Pr(C) + Pr(D) = 0 + .4 + __ + __ = ____. — .4, .2, 1

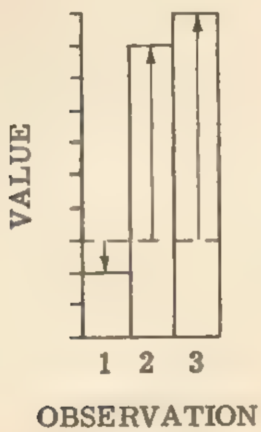
96. (Continued)

must exactly balance the negative deviations. In other words, the total length of the arrows representing the positive deviations must be exactly the ________ as the total length of the arrows representing negative deviations; this is apparently true when the reference value is 4. Thus, _____ is the **mean** of these three observations.

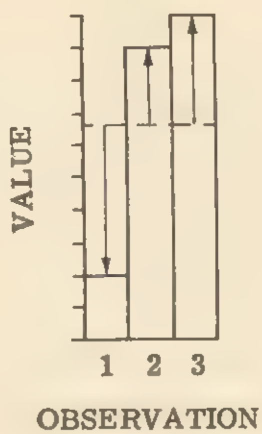
same

4

97. Consider the following two graphs.



GRAPH A



GRAPH B

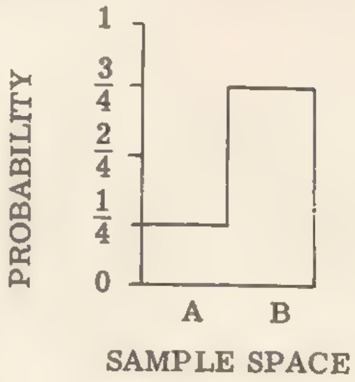
Remember, the mean is that reference value from which the sum of the deviations equals 0. Keeping this in mind, it would appear that the reference value shown in Graph ___ (A/B) was more likely the **mean** than is the reference value in the other graph.

B

98. The reference value in Graph B is more likely the mean because the two positive deviations added together almost equal the negative deviation in size, whereas the two positive deviations in Graph A added together appear to be ______________ (larger/ smaller) than the size of the negative deviation in that graph.

larger

51. Another way of representing a probability distribution is with a graph, such as the one shown below.



Notice that the **sample space** consists of two possible outcomes: ________ and ________. — A, B

52. The number assigned to each member of the sample space by the probability distribution is indicated by the height of the column over each outcome. Thus, the ________________ (number) assigned to outcome A equals $\frac{1}{4}$. — probability

53. The probability assigned to outcome B equal ______. — $\frac{3}{4}$

54. Pr(A) + Pr(B) equals ____ + ____, which equals _____, just as the total of all probabilities must in any probability distribution. — $\frac{1}{4}$, $\frac{3}{4}$, 1

99. The mean is a useful way of representing the central tendency of a distribution. However, we do not yet have a simple way of calculating the mean from an actual collection of data. Of course, you could try one reference value after another, calculating the deviations from each of these reference values. You could eventually locate a reference value from which the sum of the deviations equaled __________, which would indicate that reference value was the **mean** of the collection of data.

zero

If the collection of data consisted of only a few observations, you could probably find the mean in this manner. However, if the data consisted of many observations, the procedure outlined in the previous frame would not be practical. Therefore, we will consider a procedure whereby we can calculate the mean of any collection of data according to a simple rule. A **rule** for calculating the value of some statistic is called the **formula** for that statistic. Thus, a rule for calculating the mean would be called a ____________ for the mean.

formula

100. In order to discuss ways of calculating specific statistics, it is often useful to talk about data in general rather than about particular observed values. For example, consider the two tables of raw data shown below.

TABLE A

Observation	Value
1	20
2	10
3	42
4	52

TABLE B

Observation	Value
1	10
2	20
3	0
4	52

Each table lists the values for ____ observations of a
4/6
numerical variable. (You might wish to insert a bookmark here since we will refer to these tables in later frames.)

4

46. If the four outcomes listed in the first column of the following table represented a sample space, the missing probability in column 2 would equal ________. 1/4

Outcome	Probability
A	1/4
B	1/4
C	1/4
D	____

47. Pr(D) must equal 1/4, since the sum of the probabilities for A, B, and C equals ______ and the total of all four probabilities must equal ______. 3/4 1

48. Notice that of the two tables shown below, Table ______ (A/B) could not possibly represent a probability distribution. B

A

Outcome	Probability
A	.2
B	.2
C	.6

B

Outcome	Probability
A	.2
B	.3
C	.6

49. Table B (above) could **not** be a probability distribution since the total of the probabilities equals ______, whereas it should equal ______. 1.1 ($1 \frac{1}{10}$) 1

50. Just as a proportion can be represented by either a decimal or a fraction, a probability can be represented by either a decimal or a fraction. If a sample space consisted of two possible outcomes — A and B — and the probability of A equaled $\frac{1}{4}$, the probability of B could be represented by either the fraction ________ or by the decimal _______. $\frac{3}{4}$ .75

101. The term "**observation 1**" specifies a particular value in each table. It specifies the value 20 in Table A and the value ___ in Table B. 10

102. Similarly, the term **observation** 3 refers to the value ____ in Table A and the value ____ in Table B. 42, 0

103. Instead of writing out **observation** 1 or **observation** 2 statisticians have found it simpler to use the symbol X_1 to represent the same value as does the term **observation** 1. Thus, X_1 represents the value 20 in the previous Table A and the value ____ in Table B. 10

104. The number which appears after and just below the capital letter X in X_1 is called a **subscript**. The subscript indicates which particular observation you are representing. Thus, the subscript ______ in X_1 indicates you are representing observation 1. 1

105. Notice that the symbol X_2 has a s__________ 2 instead of a subscript 1. subscript

106. Similarly $\underset{X_2\ /\ X_3}{\underline{\qquad\qquad}}$ has a subscript 3 instead of a subscript 2. X_3

107. Since the symbol X_1 indicates the "**observation** 1," you would use the symbol X_2 to represent "**observation** ____." 2

108. Notice that the symbol X_2 can represent the $\underset{\text{first/second}}{\underline{\qquad\qquad}}$ observed value in any table of raw data, regardless of the particular value of that observation. second

41. Since Pr(A) represents the **probability** of outcome A the largest possible value of Pr(A) would be _____ and its smallest possible value would be ______. — 1, 0

42. Therefore, no matter what the actual probability distribution, Pr(A) must be a number between _____ and ______. — 1, 0

43. Imagine a sample space consisting of two outcomes — outcome A and outcome B. If Pr(A) was equal to 1, then Pr(B) must equal ______, since the sum of the probabilities must equal ______. — 0, 1

44. Just as a frequency distribution or a proportional distribution can be represented by either a table or a graph, it is possible to represent a probability distribution by either a table or a graph. For example, the table shown below represents a **probability distribution** for a sample space consisting of three possible outcomes. Notice that the three possible outcomes are listed in the ____________ (first/second) column of the table and that the numbers assigned to each of these possible outcomes are listed in the ____________ (first/second) column of the table. — first, second

Outcome	Probability
A	1/4
B	2/4
C	1/4

45. Each of the numbers in the second column is a ________________, since it is the number assigned to a member of a sample space by a probability distribution. — probability

109. In the same way, you could represent the third observation in any table of raw data by ______. The subscript_____ indicates you are referring to observation 3. X_3, 3

X_1/X_3

110. Considering the two previous tables of data, the term X_1 would refer to the value 20 in Table A and the value ____ in Table B. 10

X_3 would refer to the value ____ in Table A and the value ____ in Table B. 42 0

111. Notice how you could represent any collection of **three** observations with X_1, X_2, and X_3.

If your data consisted of X_1, X_2, X_3, and X_4, then there would be ______ observations in your collection of data. four

112. Statisticians often use a capital letter N to represent the **number of observations** in a collection of data. If the collection of data consisted of four observations, N would equal 4. If the data consisted of ten observations, N would equal _____. 10

113. Since N represents the __________ of observations in a table of raw data, you could represent any collection of N observations with a column of X's starting with X_1, then X_2, and so on, until X_N. number

This way of representing tables of data is useful because you can describe a general rule or formula for calculating some statistic without speaking of particular values. For example, you could represent the procedure for getting the **sum** or **total** of a collection of five observations with the formula:

$$\text{Total} = X_1 + X_2 + \_\_\_ + \_\_\_ + X_5$$ X_3, X_4

34. The largest probability you could assign to any member of a sample space would be ________ and the smallest would be _______.

1

0

35. Notice the similarity between numbers (probabilities) assigned to members of a sample space by a probability distribution and the proportions in a relative frequency distribution. The largest value of a **probability** is ____, and the largest value of a **proportion** is _____.

1

1

36. Similarly, the smallest value of a proportion would be ______, and the smallest value of a probability would be ______.

0

0

37. We will use the symbol Pr(A) to represent the probability of outcome A. In other words, Pr(A) represents the number assigned to outcome A, where A is a member of a _______________ _______________.

sample space

38. Similarly, if B is another member of a sample space, you can represent the number assigned to it by a probability distribution with the symbol Pr(__).

B

39. Suppose, A, B, and C represent members of a sample space consisting of 3 possible outcomes. According to our definition of a **probability distribution**,

$$\Pr(A) + \Pr(B) + \Pr(C) = \_\_\_\_\_\_ .$$

1

40. The total of the three probabilities in the previous frame would have to equal 1, since the **total** of the numbers (probabilities) assigned by a probability distribution to members of a **sample space** must equal _______.

1

114. You could write a general formula for finding the total of any collection of data consisting of three observations as follows:

Total = $X_1 + X_2 + X_3$

It would become awkward to write a formula of this sort if the collection of data consisted of very many observations. For example, if the data consisted of one hundred observations, we would have to write out a string of X values with plus signs between them, beginning with X_1 and ending with an X having the subscript _____. 100

115. One way you could simplify the formula would be to write out the phrase: **sum all the X' s**, which would mean ______________ together the values of all of the observations in the table of data. add

add/ multiply

116. Statisticians have a special shorthand way of writing the phrase: **sum all the X' s**. Instead of writing out the whole phrase, they simply write $\sum X$. Thus, if your data consisted of two observations, $\sum X$ would equal X_1 plus X_2. If your data consisted of three observations, $\sum X$ would equal ____ + ____ + ____. X_1, X_2, X_3

117. You are now in a position to write a simple formula for the total of all the observations in any collection of data. Using the shorthand way of writing "**sum all the X' s**," you could write the formula for the total as

Total = ____ X $\sum$

29. Since each of the numbers assigned to a member of a sample space is called a probability, we could describe a probability distribution as an assignment of a p____________ to every member of a sample space, such that the sum of these probabilities equals 1. — probability

30. Three different assignments to numbers to the sample space for one toss of a coin are listed below. Assignment ________ (A/ B/ C), however, could **not** be a probability distribution. — C

Assignment A:

heads - $\frac{1}{2}$, tails - $\frac{1}{2}$

Assignment B:

heads - .9, tails - .1

Assignment C:

heads - .8, tails - .4

31. The assignment of the number .8 to **heads** and number .4 to **tails** could **not** be a probability distribution since the total of the numbers __________ (does/ does not) equal 1. — does not

32. Since the total of all the numbers assigned to the members of the sample space must equal 1, the largest number you could assign to any member of a sample space would be _____. — 1

33. If one outcome in a sample space were assigned the number 1, then all the other outcomes would have to be assigned the value _____. — 0

118. The symbol $\sum$ is a capital Greek letter called **sigma**. Thus, the Greek letter _______ in the expression $\sum X$ indicates to you that the expression represents the _____ of all the X's (observations). sigma sum

119 The Greek sigma ($\sum$) is often referred to as a **summation** symbol, since $\sum X$ represents the _______ of the observed values. sum

120. Let's return now to the original task of finding a formula for the mean. Basically, this involves stating the formal relationship between the mean of the data and the observed values by a mathematical equation, and then (by simple algebra) rearranging this equation until it is in the form of a formula for the mean.

We shall begin by considering a collection of data consisting of three observations. In other words, N = ______ for this collection of data. three

121. Just as we can represent observation 1 by X_1 without indicating any particular value, we will represent the **mean** by $\bar{x}$. In this way, we can represent the mean without indicating any particular ______ for the mean. value

122. You could represent the **deviation** of observation 1 from the mean as $X_1 - \bar{x}$. In a similar way, you would represent the deviation of observation 2 from the mean as (_____ $- \bar{x}$). X_2

123. If the collection of data consisted of three observations, you could represent the sum (total) of the three deviations from the mean by $(X_1 - \bar{x}) + (X_2 - \bar{x}) +$ (___ + ___). X_3, $\bar{x}$

124. If your three observations had the values 10, 2, and 4, you could put these actual values in the previous formula and write it as $(10 - \bar{x})$ + (___ $- \bar{x}$) + (___ $- \bar{x}$). 2, 4

24. (Continued)

five possible outcomes occurred. The six possible outcomes are said to be **mutually exclusive** since the occurrence of any one outcome ex__________ the occurrence of any other possible outcomes.

excludes

25. To review, we have begun our description of Probability Theory by defining a **sample space** as a list of all __________ __________ of a process.

possible outcomes

26. Furthermore, we will suppose that **one and only one** of the outcomes can occur. That is what is meant when we said that the outcomes are m__________ __________.

mutually exclusive

27. The next basic concept of Probability Theory we will consider is that of **probability distribution**. A probability distribution is simply the assignment of a non-negative number to each member of a sample space such that the sum of all these numbers equals 1. Using the sample space for one toss of a coin as an example, therefore, we could assign the number $\frac{1}{2}$ to the outcome **heads** and the number $\frac{1}{2}$ to the outcome **tails**. Each of the outcomes in the sample space has been assigned a number such that the total of these numbers equals 1. This assignment of numbers to members of a sample space is an example of a p__________ distribution.

probability

28. Each of the numbers assigned to a member of a sample space is called a **probability**. Therefore, the number $\frac{1}{2}$ assigned to the outcome heads in the previous illustration is a __________.

probability

125. According to our previous discussion of the mean, you know the sum of the deviations from the mean must equal _____. Therefore, the following formula ______________ (could/could not) be true if $\bar{x}$ is the mean of these three observations: — zero, could not

$$(10 - \bar{x}) + (2 - \bar{x}) + (4 - \bar{x}) = 5$$

126. If your data consisted of three observations and the following equation were true, 10 would be the _______ of these three observations. — mean

$$(X_1 - 10) + (X_2 - 10) + (X_3 - 10) = 0$$

127. In other words, without actually stating the values of the observations nor the actual value of the mean, we know (from the definition of the mean) that the following equation is ____________ for any collection of three observations. — true

$$(X_1 - \bar{x}) + (X_2 - \bar{x}) + (X_3 - \bar{x}) = 0$$

128. According to simple algebra, we ______________ (could/could not) remove the parentheses in the previous equation and write it as follows: — could

$$X_1 - \bar{x} + X_2 - \bar{x} + X_3 - \bar{x} = 0$$

129. Furthermore, we could rearrange the symbols on the left-hand side of the previous equation so that the equation reads as follows: — X_3, $\bar{x}$

$$X_1 + X_2 + \_\_\_ - \bar{x} - \bar{x} - \_\_\_ = 0$$

21. While these games of chance are **examples** of random processes, a random process can be defined in general as **any** process that can be represented by ____________ ________________.

Probability Theory

22. A complete introduction to the mathematical theory of probability is far beyond the scope of this program. It will be possible, however, to illustrate the character of Probability Theory and to introduce some of its more fundamental concepts. One of the basic concepts involved in Probability Theory is the idea of a **sample space**. Whenever a process can have several possible outcomes, a list of all possible outcomes can be thought of as a **sample space**. For example, when you toss a coin the two possible outcomes are heads and tails. (We will ignore the possibility of its landing on its edge for this illustration.) Thus, a **sample space** for the simple process of tossing a coin is represented by the following list of possible outcomes: ___________ and ___________.

heads, tails

23. A **sample space** for a single roll of the die consists of a list of all possible outcomes. In other words, the list 1, 2, 3, 4, 5, and 6 represents the _________ ________ for a single roll of a die, where each of the six numbers is a possible number of dots showing after each roll.

sample space

24. Notice that **only one** of the possible outcomes in the sample space can occur each time the die is rolled. In this sense, the occurrence of a particular outcome **excludes** the occurrence of any of the other outcomes. Thus, to say that a die was rolled and that one dot was showing is equivalent to saying that none of the other

130. Remember, the left-hand side of the previous equation is the **sum of the deviations from the mean**. Notice how this sum is made up of the total of all the scores from which we subtract the mean $\underset{\text{three/four}}{\underline{\qquad\qquad}}$ times. three

131. We can represent the total of the three observations by $\sum X$. Also, subtracting the mean three times is the same as subtracting $3\bar{x}$. Therefore, we could write the previous equation as $\sum X - \_\_\_\_ \bar{x} = 0$. 3

132. The previous equation says that when we subtract $3\bar{x}$ from the sum of the three observations, our answer is zero. Therefore, the total of the three observations ($\sum X$) must be exactly equal to 3 _____. $\bar{x}$

133. We started with an equation that said the sum of the deviations from the mean equals zero, and we have proceeded to the following equation:

$$\sum X = 3\bar{x}$$

If we divide both sides of the previous equation by 3, the result would be

$$\frac{\sum X}{3} = \frac{3\bar{x}}{3}$$

The 3's on the right-hand side of the equation cancel each other out, leaving us with the equation:

$$\frac{\sum X}{3} = \_\_\_\_$$

$\bar{x}$

18. Probability theory can be defined simply as a mathematical way of presenting __________ processes. In fact, mathematicians originally developed probability as a way of representing the random process in a game of chance. Later, they found Probability Theory was a useful way of representing processes other than in games of chance. It became useful to think of a random process as **any** process that could be represented by Probability Theory. Therefore, it is possible to say that any process that can be represented by Probability Theory is a ______________ process.

random

random

19. As a psychologist, you will be interested in Probability Theory primarily because it is a way of representing the uncertainty involved in random sampling procedures. It is even possible to use Probability Theory to predict the form of sampling distributions without actually collecting the samples necessary to form an **experimental** sampling distribution. We noted earlier that an experimental sampling distribution is based on an actual collection of samples, while a **theoretical** sampling distribution is **not** based on actual samples but rather on logical or theoretical considerations. Thus, the sampling distribution predicted by means of Probability Theory would be a(n) ____________________ sampling
experimental/theoretical
distribution.

theoretical

20. Remember, Probability Theory was originally devised as a mathematical way of representing g__________ of c__________, such as dice and roulette.

games

chance

134. This equation is a **formula** for the mean, since it says that if we added together our three values and divided this sum by the number of observations (3), the result would equal the mean.

Let's see if this formula works for a particular example. Suppose your data consisted of the values 2, 8, and 2. The formula says to first add the observations. This would give you a sum of _____. Dividing this sum by 3 would give you a result of _____.

12
4

135. In other words,

$$\frac{\sum X}{3} = \frac{6 + 2 + 2}{3} = \frac{12}{3} = 4$$

Therefore, according to the formula, the **mean** of the three observations is _____. We can check to see whether 4 really is the mean of the previous observations by considering the following table:

4

OBSERVATION	VALUE	DEVIATION FROM 4
1	8	4
2	2	-2
3	2	-2

The sum of the deviations from the reference value 4 equals _____. Therefore, we know that 4 is the _____ of the 3 observations.

zero, mean

13. Each time you roll a die, there are _______ possible outcomes since the die can fall with either one, two, three, four, five, or six dots showing on its face. — six

14. While you are uncertain exactly how many dots will show on any particular roll of the die, you can describe this uncertainty. For example, you would expect each of the possible outcomes to occur about as often as the other possible outcomes. Thus, you might make the following kind of confidence statement:

> "I would expect the die to fall with one dot showing about _______ times out of every 600 tosses." — 100

15. The answer to the previous frame was 100 since each of the six possible outcomes would be expected to occur about as often as any other. Since there are _____ possible outcomes, each outcome would occur **about** 100 times in every 600 tosses. — 6

16. Another example of a **random process** is the kind of **random sampling procedure** we described earlier. According to this sampling procedure, each member of a population has the ____________ opportunity to be included in a sample as has any other member. — same

17. In this section, we will consider a type of mathematics called **Probability Theory**, which is useful in describing random processes. In other words, this type of mathematics ________________ (would/would not) be useful in describing games of chance. — would

136. The formula we derived was for a collection of **three** observations. Suppose the data consisted of 4 observations. We could write the following equation, which specifies the relationship between the mean and those 4 observations:

$(X_1 - \bar{x}) + (X_2 - \bar{x}) + (X_3 - \bar{x}) + (X_4 - \bar{x}) =$ _____. 0

137. We could remove the parentheses in the previous equation and rearrange the terms so that the equation reads as follows:

$X_1 + X_2 + X_3 + X_4 - \bar{x} -$ ____ $-$ ____ $-$ ____ $= 0$ $\bar{x}$, $\bar{x}$, $\bar{x}$

138. Using our shorthand way of writing this, we could say that:

_____ $X - 4\bar{x} = 0$ $\sum$

139. Remember, the sum of the deviations from the mean equals $\sum X - 3\bar{x}$ when the data consists of 3 observations, and $\sum X - 4\bar{x}$ when the data consists of 4 observations. No matter how many observations there are in the collection of data, it is not hard to see that the following equation __________ (would/would not) be true. would

$$\sum X - N\bar{x} = 0$$

where N is the number of observations.

140. Just as we did earlier, we could rearrange this equation to read: $\sum X =$ ____ $\bar{x}$. Then, dividing both sides of the equation by N, we would finally arrive at the following formula for the mean: N

$$\frac{\sum x}{N} = \frac{N\bar{x}}{N} = \_\_\_\_$$ $\bar{x}$

9. You are often uncertain about the outcome of many processes. For example, when you roll a die, you are ______________ (certain/uncertain) regarding the outcome. When you draw a random sample from a population, you are ______________ (certain/uncertain) about the particular sample you will obtain.

uncertain

uncertain

10. A process in which there is uncertainty concerning outcomes can be described as a **random process**. The most common examples of a random process are games of chance, such as dice or roulette. It is the u__________ of how many dots will show when the die is thrown and the u______________ over where the ball will come to rest in roulette that makes these games **random processes**.

uncertainty

uncertainty

11. A simple game of chance is one in which a person bets which face will show when a coin is tossed. A simple game of this sort would also be a r______________ p______________ because it is uncertain which face will show on a given toss.

random process

12. While you would be uncertain about the outcome of tossing a coin, it is possible to say something about your uncertainty in this kind of random process. For example, if the coin were a normal one, tossed in a normal manner, you would expect heads to occur about as often as you would tails. In other words, you might make the following kind of confidence statement concerning the outcome of any particular toss.

"I would expect to obtain about _____ heads out of every 100 tosses of the coin."

50

140. (Continued)

This formula says the mean of a collection of N observations is equal to the total of all of the observations divided by the ______ of those observations. number

141. Thus, if you had a collection of 10 observations and the total of all the observations were 50, the mean of the observations would equal _____ divided by _____. In other words, the mean would equal _____. 50, 10 5

142. Let's test this formula on the following collection of 5 observations.

OBSERVATION	VALUE	DEVIATIONS FROM 6
1	10	4
2	6	0
3	6	0
4	6	0
5	2	-4

Notice that $\sum X$ = ______ for this data (because $\sum X$ is a shorthand way of writing "sum of all the values"). 30

143. $\sum X$ equals 30, and N equals _____; thus, our formula for the mean says to divide _____ by _____, which would give a result of _____. 5 30, 5 6

144. The formula says 6 is the mean of these five observations. In the third column of the table, we have listed the deviations of each observation from the reference value _____. It is clear that the sum of these deviations equals _____. This indicates that 6 __________ (is/is not) the mean of these 5 observations. 6 0, is

8. In everyday conversation, it is often sufficient to say that one thing is more likely than another in order to express different degrees of uncertainty. As a scientist, however, it is often important to describe your **degree of uncertainty** more precisely. The **confidence statements** we considered earlier are examples of a more precise way of describing uncertainty. For example, both of the following confidence statements imply uncertainty that the mean of a sample drawn from a population in a particular manner will be within one standard deviation of the population mean:

Statement A

> Nine-tenths of all samples obtained in this manner had means within one standard deviation of the population mean.

Statement B

> Six-tenths of all the samples obtained in this manner had sample means within one standard deviation of the population mean.

Although both statements imply uncertainty, you would be **more** certain of obtaining a sample whose mean was within one standard deviation of the population mean if Statement ________ (A/B) were true. A

You would be **more** certain of obtaining a sample mean within one standard deviation of the true mean if Statement A were true, since, according to Statement A, ________ (.9/.6) of the samples obtained in that manner had sample means within one standard deviation of the true mean, whereas according to Statement B, only ________ (.6/.9) had means within one σ of μ. .9 .6

145. We have considered three ways of representing the central central tendency of a distribution: the m______, the m________ and the m_________. — mode; median, mean

146. Often a distribution will have a different mean, median, and mode. The data shown in the previous table, however, has a mode of ____, a median of ______, and a mean of ____. — 6, 6; 6

147. Six is the __________ of the previous collection of data since this is the most frequently occurring value. Six is also the ___________ of the previous collection of data since there is, in the collection of data, one value greater than six and one value less than six. Finally, 6 is the __________ of the previous collection of data because the sum of the deviations from 6 equals zero. — mode; median; mean

148. The mean is a useful way of representing the central tendency of a distribution because its value depends on every value in the distribution. The value of the _________________ (mean/ median/ mode) is determined only on the basis of the most frequently occurring value in the collection of data. Finally, while you know that there are the same number of values above the _________________ (mean/ median/ mode) as there are below it, you _________ (do/ do not) know how far above or how far below the median these values are. Thus, each of the various ways of representing the central tendency of a distribution has its own peculiar features. — mode; median; do not

4. Almost any event you speak of **as if** it were a certainty will be revealed to have some uncertainty when examined closely in this way. Naturally, there are many events which are not spoken of as if they were certainties. For example, you might say, "It is **very likely** the Yankees will win the World Series this year." This statement implies that you __________ certain the
are/are not
Yankees will win the World Series. — are not

5. You might tell someone that you will probably be home this evening and the word **probably** implies you __________ absolutely certain whether you will be home
are/are not
or not. — are not

6. It is possible for a statement to imply more or less uncertainty. For example, both the statement "I will **probably** be home tonight" and the statement "I **might** be home tonight" imply that you are ______________
certain/uncertain
about being home. — uncertain

7. While both statements imply some uncertainty, the statement "I will probably be home tonight" implies that you are __________ certain of being home than does the
more/less
statement "I might be home tonight." — more

1. We have just seen that the mean, median, and mode are three ________________ that characterize the ________________ ________________ of a distribution. — statistics; central tendency

2. Besides their central tendency, there is another important characteristic of distributions, one that is **not** represented by either the mean, the median or the mode. Some collections of data are composed of many similar values, while in other distributions the values might vary considerably. For example, the values in Data $\underset{A/B}{\underline{\qquad}}$ (below) are all very similar to each other, while the values in Data $\underset{A/B}{\underline{\qquad}}$ are much more widely separated. — A; B

Data A: 20, 21, 20, 19, 20

Data B: 2, 38, 20, 5, 35

3. We have listed the two previous collections of data in the following tables:

Observation	Value	Deviation From 20
1	20	0
2	21	1
3	20	0
4	19	-1
5	20	0

TABLE A

Section VII: **Probability and Theoretical Sampling Distributions**

1. You sometimes speak **as if** an event in the future could be predicted with **certainty**. For example, you might say, "When I press this button on the wall, the light on the ceiling will turn on." This statement implies you are ________________ (certain/uncertain) the light will turn on when you press the switch. — certain

2. Naturally, there is always the possibility that the switch may have been broken or that the light bulb burned out since the last time you pressed the switch. If the switch has almost always worked in the past, however, it is often simpler to speak a____ i___ you were certain that the light would go on, even though there is some possibility that it will not. — as if

3. If you were explaining what would happen when you tossed a coin, you might say it will land either **heads** or **tails**, **as if** it were certain that the coin would not land balanced on one edge and, therefore, be neither heads nor tails. Once again, it is simpler to speak **as if** you were certain the coin would land either heads or tails, since the proportion of times it would land balanced over a very large number of tosses is so____________ (small/large) that the possibility can be ignored for most purposes. — small

3. (Continued)

Observation	Value	Deviation From 20
1	2	-18
2	38	+18
3	20	0
4	5	-15
5	35	+15

TABLE B

Notice that the tables indicate the deviation of each value from the reference value _____. 20

4. The sum of the deviations in Table A is _____. 0

Therefore, the mean of Data A is _____. 20

5. The sum of the deviations in Table B is _____; therefore, the mean of Data B and the mean of Data A are 0

_______________. the same

the same/different

6. While the means of both distributions are identical, there is an interesting difference in the two distributions. If we ignore whether or not a deviation is positive or negative and only consider its absolute size, it is clear that the deviations in Table _______ tend to be larger B

A / B

than the deviations in the other table. In other words, the values in Table _______ are more spread out B

A / B

(dispersed) around the mean value of 20.

6. If a population consists of a specific number of observations, it is called:

   a. an infinite population.

   b. an unlimited population.

   c. a disconnected population.

   d. none of the above ________ d

TRUE OR FALSE:

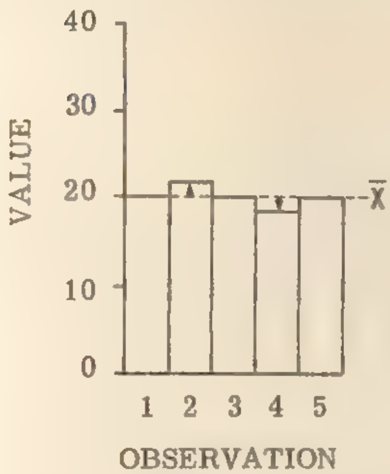
7. If the true mean of a population were 27 and your estimate were 25, your absolute error of estimate would be 2. ________ true

8. Sampling procedure in which all members of the population have the same opportunity of being selected is called a random sampling procedure. ________ true

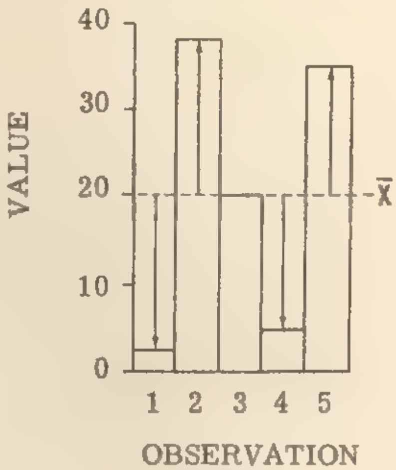
9. The larger the size of the samples, the more frequently you would expect the sample mean to be far from the population mean. ________ false

10. The risk of obtaining a highly unrepresentative sample increases with the size of the sample. ________ false

7. The manner in which the deviations from the mean are different in the two previous collections of data can be illustrated by the following graphs of raw data.



GRAPH A



GRAPH B

We have represented deviations from the mean with arrows, as we have done previously. If we ignore the direction in which an arrow is pointing and only consider its length, it is clear that the deviations from the mean in Data ______ (A / B) are larger than those of the other collection. B

8. Notice how the heights of the columns in Graph ______ (A / B) are very similar, whereas the heights of the columns in the other graph tend to change or **vary** much more from one observation to another. A

9. Since the height of a column represents the value of that observation, we could say that the values on Graph ______ (A / B) change or **vary** more than do those on the other graph. B

REVIEW III

FILL IN THE BLANKS:

1. Distributions in which most of the scores are piled up near one end are called ______________. skewed

2. If a distribution has most of its observations piled up near the high values, it is said to be ______________ ______________. negatively skewed

3. If a distribution has a standard deviation of 3 and a mean of 10, then a score of 16 would represent a standard score of ________. 2

MULTIPLE CHOICE:

4. A value identical to the mean has a standard score of:

   a. 0.

   b. 1.

   c. 2.

   d. none of the above ________ a

5. The sample, as used by statisticians, refers to:

   a. a very small, but complete, collection of data.

   b. a collection of data which is considered as part of a larger, complete collection of data.

   c. bell-shaped.

   d. none of the above ________ b

10. The more the values in a collection of data vary, the more **variability** that collection of data is said to have. Thus, we would say that the **variability** of Data ______ was greater than the variability of the other collection of data. B

A / B

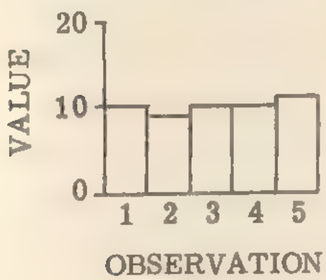
11. Data composed of many widely different values is said to have a great deal of **variability**. On the other hand, a collection of data in which the values are very similar or close together could be described as having little **variability**. Therefore, of the following two collections of data, Data ______ would be described as having more **variability** than the other collection. B

A / B

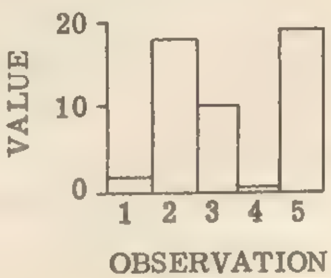
Data A: 10, 9, 10, 10, 11

Data B: 2, 18, 10, 1, 19

12. We could illustrate the difference in **variability** of the two previous distributions with the following graphs:



GRAPH A



GRAPH B

Earlier we defined a variable as something that changed or varied. The more something changes or varies, the more **variable** it is said to be. Thus, the observed values are more variable in Graph ______ than in the other graph. B

A / B

263. (Continued)

this **risk**. It is important to bear in mind that the sampling distribution is determined **both** by the characteristics of the population you are sampling and the **procedure** by which you have obtained the samples. The experimental sampling distributions we have just considered are based on a ________________ sampling procedure.

random

264. In the earlier illustration,in which samples of size 10 were drawn from the population of 10, 000 college students, we were interested in student opinion on a particular question. Suppose the same procedure were followed to obtain 100 samples from the same students but on an entirely different question. If the population proportion of favorable opinions on this question were not the same as the population proportion of favorable opinions on the previous question, you would expect the experimental sampling distribution to be ________________.
similar/different

similar

265. Suppose you obtained 100 samples of opinions using a **biased sampling** procedure. In this case, you ________________ know whether to expect the new
would/would not
experimental sampling distribution to be similar to the old experimental sampling distribution. In the next section, we will consider the manner in which you can predict the distribution of sample statistics obtained by a **random** sampling procedure without actually collecting the samples. In other words, we will consider how to determine a ________________ sampling
theoretical/experimental
distribution.

would not

theoretical

13. The more variable the values in a collection of data, the more **variability** that collection is said to have. In other words, the more the values in a collection of data change or vary from observation to observation the more ______________ that collection is said to have. variability

14. If all the values in a collection of data were very similar, the data would **not** have much variability. If all the values are very similar, you would expect the difference between the largest value and the smallest value to be __________.
(large/small) small

In other words, if the difference between the largest and smallest value in a collection of data were very small, the data ______________ have much variability.
(would/would not) would not

15. We often use the **difference** between the largest and smallest value in a collection of data as a statistic representing the variability of that data. We call this statistic the **range**. In other words, you could represent the variability of a collection of data by finding the difference between the largest and the smallest value in that collection of data. This difference is a statistic called the ____________. range

16. If your data consisted of the values 100, 50, 10, 75, and 80, then the __________ of these values would be 90, since this value is the difference between the highest value (100) and the lowest observed value (10). range

260. (Continued)

On the other hand, every banana used in your sample would be unsuitable for sale. This is an example of what is called **destructive** sampling, since the bananas included in each sample are **destroyed**. To determine the size of your sample, you would have to compare the ________ (gain/loss) associated with having a large and more likely representative sample to the ________ (gain/loss) produced by destroying those members of the population that would be included in the sample.

gain

loss

261. A tire manufacturer submitted **some** of the tires produced at his factory to pressure until they exploded. In this way he determined the maximum pressure he could expect tires produced at his factory to withstand. This would be another example of d____________ sampling.

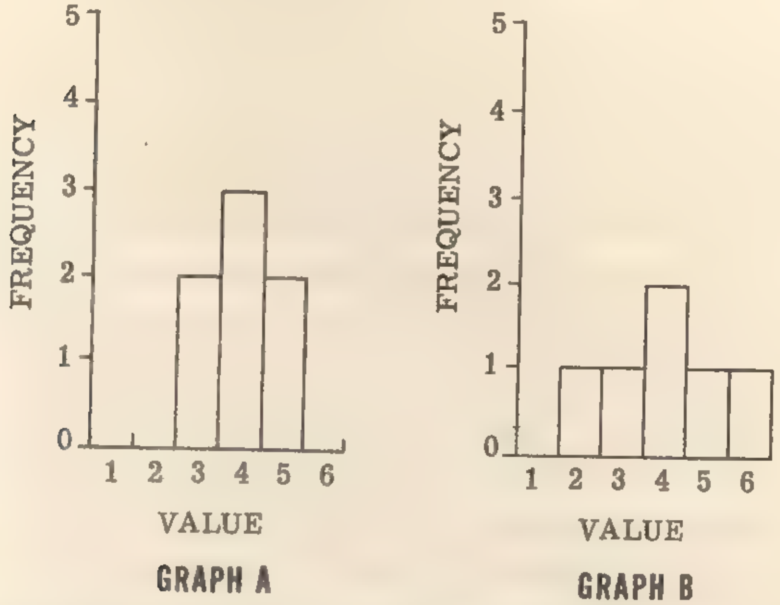
destructive

262. Sometimes it is impossible to obtain a complete collection of data. In such a case you would have to settle for a sample. For example, suppose your population consisted of the heights of all the people on the North American continent 100 years before the arrival of Columbus. If you only knew the heights of three skeletons found by archeologists and identified as belonging to that population, you would only have a ____________ from the ____________.

sample, population

263. While there are many advantages to obtaining samples, you always run the **risk** of obtaining an unrepresentative sample. You have seen how knowledge of the distribution of sample statistics can help you in specifying

17. The relation between the range and what we have referred to as the variability of a distribution can be illustrated on the following frequency distribution graphs:



The data shown on both of these graphs have the same mean, but the data on Graph B is more **variable** than the data on the other graph.

Notice that the **difference** between the highest and lowest observed values on Graph A equals ____ minus ______. 5, 3

The difference between the highest and lowest value on Graph B equals _____ minus ______. 6, 2

18. Since the range of a collection of data is the difference between the highest and lowest values, the **range** of the data on Graph A is _____ and on Graph B, _______. 2, 4

258. It is quite apparent that the frequencies of sample means whose absolute difference from the population mean was 1 or less would ________________ as the size
increase/decrease
of the sample was increased. increase

259. In this section of the program, we have been considering the problem of characterizing a population on the basis of a sample. There are many advantages to simply obtaining a sample rather than a complete collection of data on the population. For example, it might be very **difficult** to obtain a complete collection of data. If you wanted to know something about every person in the United States, you could do one of two things. You could conduct a national census (an extremely difficult task, which the government only attempts to undertake every few years), or you could obtain only a ______________ from this population, rather than obtaining a complete collection of data. sample

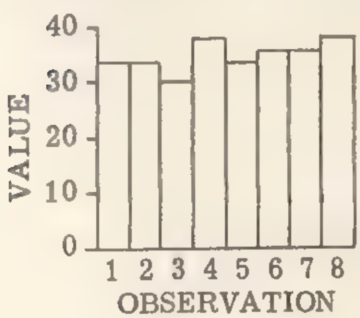
260. Suppose you were a merchant who had just received a shipment of 50, 000 bananas from South America. You could determine the condition of these bananas by peeling and tasting each of them. This is obviously not a very sensible proposition since a merchant could not market 50, 000 peeled and nibbled bananas. On the other hand, you might answer your question satisfactorily by merely taking a sample from this population. You could peel and taste only some of the 50, 000 bananas. You would have some incentive for making your sample large since a large sample is more likely to be representative of the population.

19. Notice that the distribution in which the values are **spread** farthest from the mean has the largest range. In other words, the distribution which is most v______________ would tend to have the largest range. variable

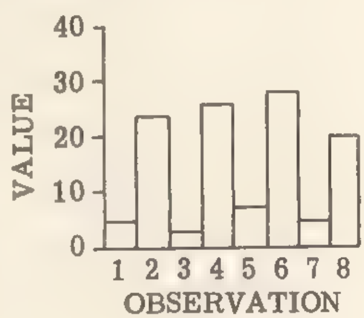
20. Suppose the largest value in a collection of data were 100 and the smallest value were 10. The range of that collection of data would be _____, since this equals ____ minus 10. 90, 100

21. Instead of considering frequency distribution graphs, let's consider how a graph of raw data indicates the range. The **range** of a collection of data is immediately obvious in a graph of **raw data**, since the range is the __________ in height of the column representing the smallest observed value and the height of the column representing the largest observed value. difference

22. The variability of a collection of data does **not** depend on the size of the values in the collection. It only depends on the **differences** in the sizes of the values. For example, consider the two graphs of **raw data** shown below. (Do not confuse these graphs with frequency graphs.)



GRAPH A



GRAPH B

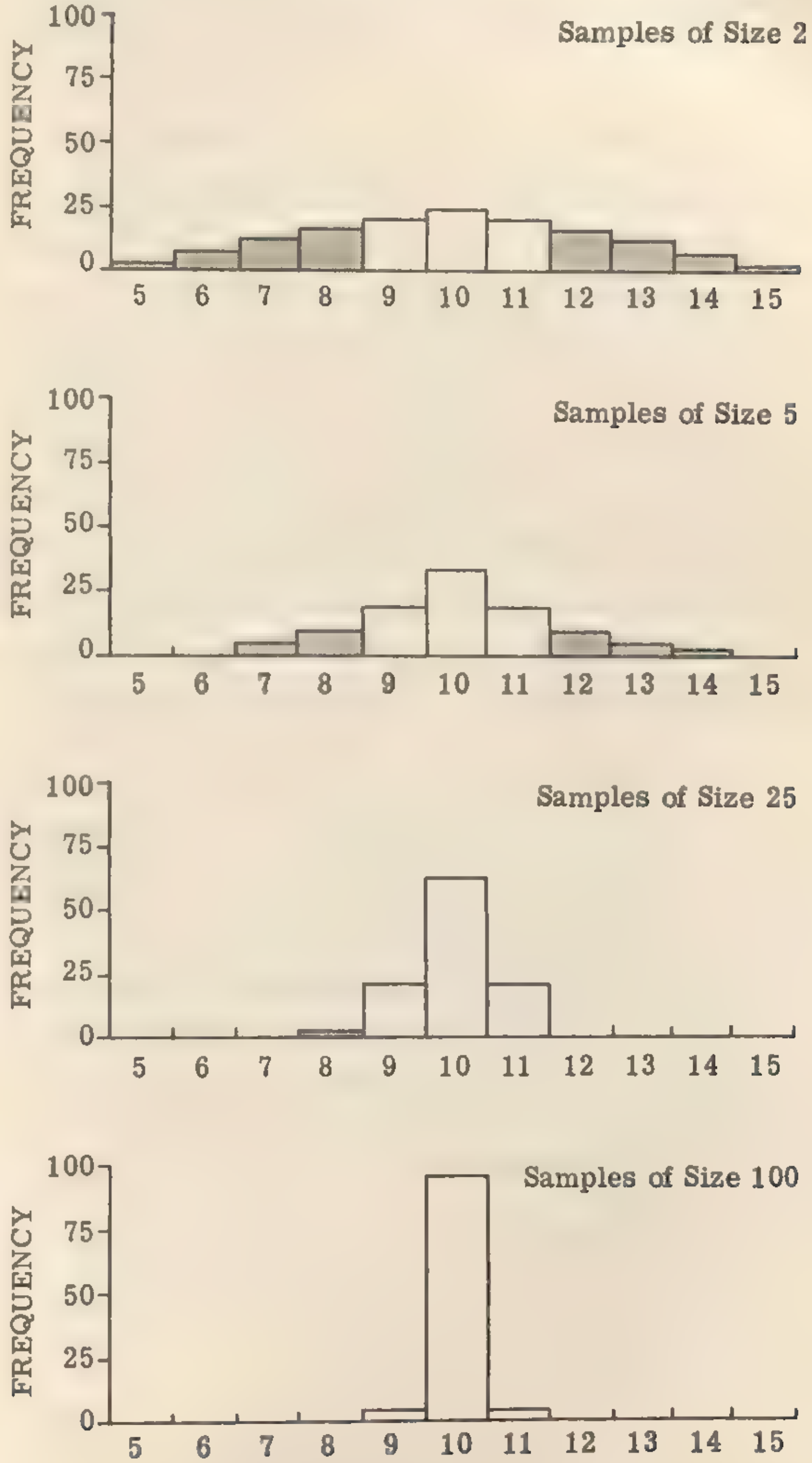
Notice that the values on Graph $\underset{A / B}{\underline{\quad\quad}}$ are all larger than those on the other graph. However, the variability of the data on Graph $\underset{A / B}{\underline{\quad\quad}}$ is greater than the variability of the other collection of data.

A

B

257. (Continued)

Frequency distributions of 100 sample means based on samples drawn at random from a large population. The population has an approximately normal distribution with a mean of 10 and a variance of 8.

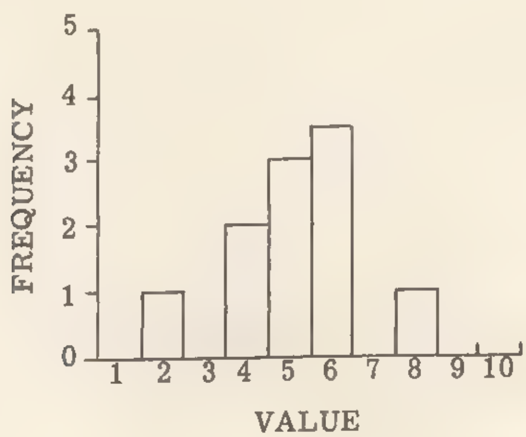


23. The range of the data in Graph ______ (A / B) would be found by subtracting the value of observation 3 from the value of observation 4. A

24. The **range** of a collection of data can also be determined from a **frequency** table by finding the difference between the smallest value having a frequency greater than 0 and the largest value having a frequency greater than 0. In the following frequency table, for example, the largest value having a frequency greater than 0 is _____ and the 12
smallest value having a frequency greater than 0 is ____. 10
Therefore, the range of this distribution is ______. The 2
value 13 was a possible value which ______ (was/was not) observed. This is why we only consider values with frequencies greater than zero. was not

Value	Frequency
10	2
11	20
12	3
13	0

25. The **range** of the distribution shown in the following frequency distribution graph is ____, since the smallest 6
observed value is ____ and the largest observed value 2
is _____. 8



255. You might think about the change in the variability of sample means as the size of the sample is increased in the following manner. If the sample is very large, you would expect to obtain many values in the center of the distribution as well as some values from each tail. The sample mean would tend, therefore, to be quite close to the true population mean. On the other hand, if the sample were very small (for example, a sample of 2), you could more easily obtain both values in your sample from the same tail of the population distribution. In this case, the sample means would be quite

______________ the population mean. different fro
similar to/different from

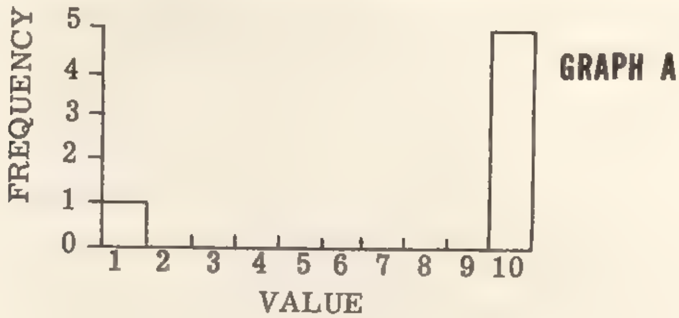
256. In the case of small samples, you run a greater risk of obtaining values from one extreme of the distribution. This would give you a highly unrepresentative picture of the population. On the other hand, the larger the size of your sample, the more likely it is that the sample will contain values from all parts of the distribution and would therefore be ______________ of the representati
representative/unrepresentative
population.

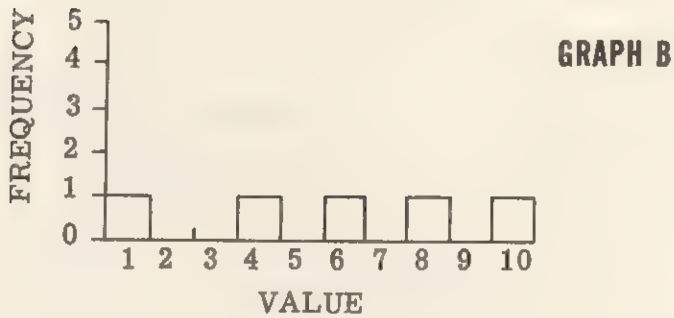
257. The following figure contains the four

______________ sampling distributions we experimenta
theoretical/experimental

considered a moment ago. However, we have shaded all the columns representing frequencies of sample means which would have led to an absolute error of estimation greater than ______. The unshaded columns represent 1
the frequencies of samples in each distribution which would have led to an error of estimation of _____ or less. 1

26. The range can sometimes give a misleading picture of the variability of a distribution. For example, consider the following two frequency distribution graphs.





Notice that the range of distribution in Graph A is ____ 9
and the range of distribution in Graph B is ______. 9

Notice that all the values except one are identical in
Graph ______ (A / B), whereas the values in the other graph were A
all different. Therefore, even though both distributions
have the **same** range, the distribution in Graph ______ (A/B) A
appears to be more variable than does the other distribution.

The problem in representing the variability of a distribution with the range is that it is determined solely by the **largest** and **smallest** observed values.

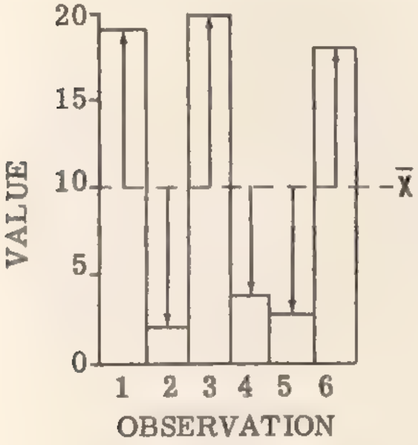
252. The value of μ (in other words, the population ________) is ______. You could say the **central tendency** of all four experimental sampling distributions __________ (was/was not) approximately 10. — mean; 10; was

253. While the experimental sampling distributions had similar central tendencies, the **variability** of the sampling distributions ______________ (increased/decreased) as the size of the sample was increased. — decreased

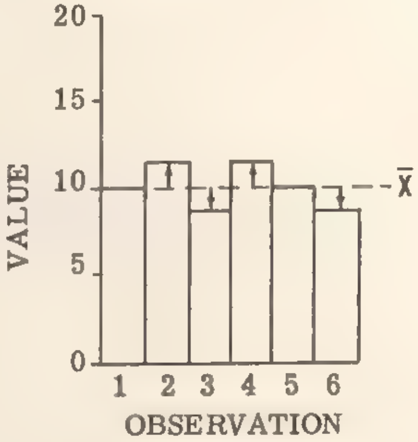
254. Notice that all four experimental sampling distributions are approximately symmetrical and somewhat bell-shaped. Since the population has been described as having a distribution that was approximately normal, you know that the population distribution was ________________ (symmetrical/asymmetrical) and had that special form of a bell-shaped distribution in which the proportion of values within any number of standard deviations was described by the graph we considered earlier. — symmetrical

Most of the values in a **normal distribution** are grouped about the mean with progressively fewer values at greater distances from the mean. In other words, the values in the tails of a normal distribution have the __________ (lowest/highest) frequency. — lowest

27. Earlier, we pointed out how the variability of a collection of data is related to the deviations of the values from their mean. For example, consider the two graphs of raw data shown below:



GRAPH A



GRAPH B

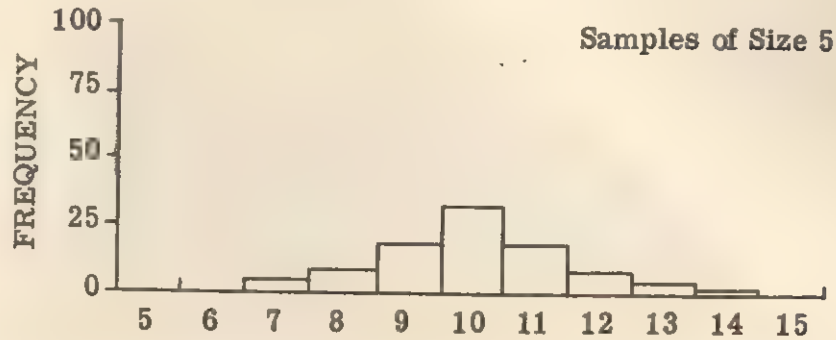
We have shown deviations from the **mean** value on both of these graphs. Thus, the mean of both distributions is ________________ (the same/different) since it equals 10 on Graph A and 10 on Graph B. the same

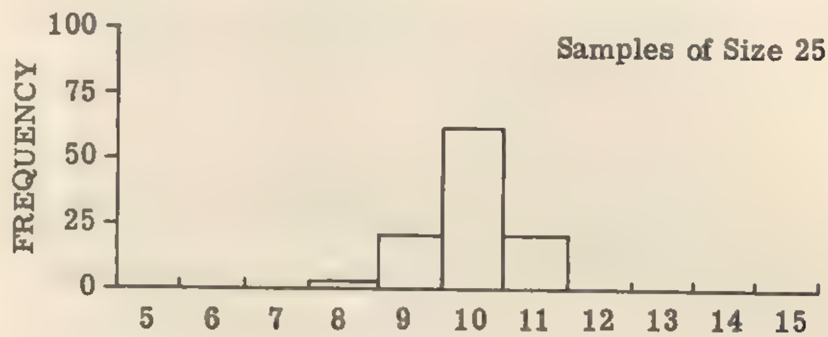
While we know that the sum of the deviations from 10 equals ______ in both distributions, there is, nevertheless, an important difference between the two distributions. 0

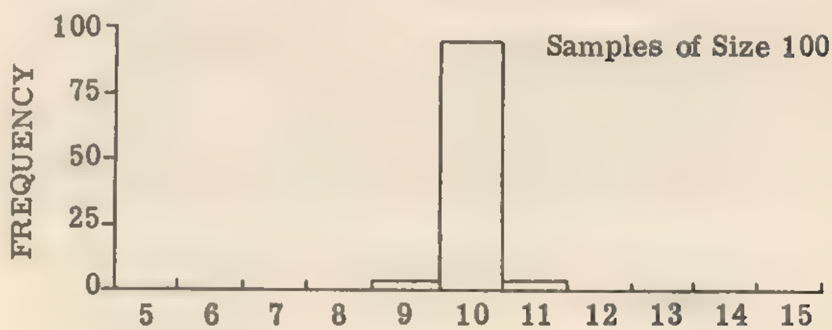
28. It is apparent that the deviations from 10 tend to be larger for the data shown on Graph ______ (A / B) than for the data in the other graph. A

29. You could describe the differences between the two distributions by saying that the value of each observation tended to vary or change more on Graph _____ (A / B) than on the other graph. A

251. (Continued)







Notice that these four frequency distribution graphs include one graph for samples of size 2, another for sample means from samples of size 5, another for sample means from samples of size 25, and the last (bottom graph) for sample means from samples of size ______. 100

30. Data containing many widely separated values of a variable is said to have more **variability** than data containing very similar values. Thus, of the two previous distributions, the Distribution in Graph ______ (A/B) could be described as having the most **variability**. — A

31. In other words, we could say that the two distributions shown in Frame 27 are similar in terms of their ______________ (central tendencies/ variability) and different in terms of their ______________ (central tendencies/ variability). — central tendencies; variability

32. The relationship between the variability of a collection of data and the size of the deviations from the mean is illustrated by the following two tables.

Observation	Value	Deviation from 10
1	15	5
2	8	-2
3	12	2
4	5	-5

TABLE A

Observation	Value	Deviation from 10
1	9	-1
2	10	0
3	11	1
4	10	0

TABLE B

Notice that the deviation of observation 4 from the mean in Table A was obtained by subtracting ____ from ____, to yield an answer of ______. — 10, 5; -5

249. In other words, you would be ________ **confident** that your absolute error of estimate would be 1 or less if the sample size were 5 than if the sample size were 25. less

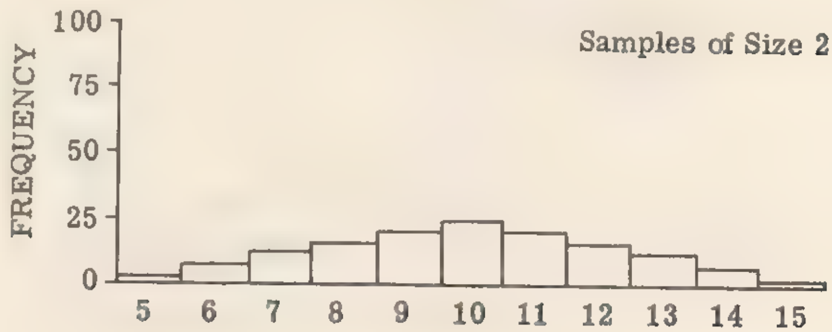
more/less

250. In general, then, as you increase the size of your sample, it appears that you can be ________ confident that the sample mean will be close to the population mean in value. more

more/less

251. The previous illustration is even more apparent when you consider the following graphs representation of the four experimental sampling distributions we have just considered.

Frequency distributions of 100 sample means based on samples drawn at random from a large population. The population has an approximately normal distribution with a mean of 10 and variance of 8.



33. Notice that the sum of the deviations in Table A equals
_____ and the sum of the deviations in Table B equals 0
_____. This indicates that the _______ of both 0, mean
distributions is 10.

34. It is clear that the values tend to be farther away from
the mean in Table _______ than they do in the other table. A
A / B

Without considering the deviations from the mean, it is
apparent that the value of the observations changed or
varied more in Table _______ than they did in the other A
A / B
table. Therefore, the _______________ of the data variability
in Table A appears to be greater than the variability of
the data in the other table.

35. We could represent the difference in variability by
calculating the **range** of each collection of data. The
range of the data in Table A would be _____, since this 10
equals ______ minus ______. 15, 5

36. Since the range of the data in Table B is _____, the 2
range of the data in Table A is ______________ than the greater
greater/ smaller
range of the data in Table B.

37. The range is not the only statistic you can use to represent
the variability of a distribution. There is another way
of characterizing the difference in variability of the
two previous collections of data. Notice that if we ignore
whether a deviation is positive or negative, the
deviations in Table ______ tend to be larger than those in A
A / B
the other table.

245. We might put this another way. We might say: Of the 100 samples of size 100, there were ________ samples which would have led to an absolute error of estimation greater than 1. no

246. We have compared all four experimental sampling distributions in terms of the number of sample means which differed from the true population value (10) by 1 or less (in other words, sample means of 9, 10, or 11). In summary, there were **45** samples of **size 2** which would have led to an absolute error of estimation of one or less. There were **76** such sample means for samples of **size 5**. There were **99** such samples of **size 25**, and **100** such samples of **size 100**.

247. According to the preceding statement, you could make the following **confidence statement**: "If I obtained samples of size ________ from this population by this random sampling procedure, I would expect the absolute error between my sample mean and the population mean to be 1 or less about 66 times in every 100 samples." (Once again, we shall assume that the 100 samples of each size are reasonably representative of the distribution of samples you would obtain through additional sampling.) 5

248. If the previous **confidence statement** had specified a sample of size 25 instead of **size 5**, you could have said: I would expect the absolute difference between the sample mean and the population mean to be one or less in about ______ out of every 100 samples. 99

37. (Continued)

You can think of the **variability** of a collection of data as the degree to which the values are spread out from the mean. In other words, if all of the values in a collection of data are very similar, they will all be clustered very close to the mean and the data will **not** have much variability. If the values in a collection of data are widely dispersed (spread out), the deviations from the mean will tend to be very ___________ and the
large/small

large

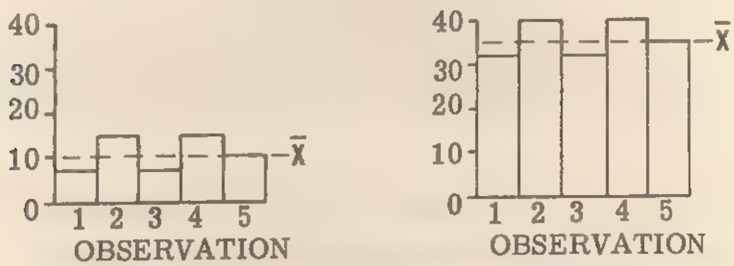
distribution could be described as having a great deal of _______________.

variability

38. We have seen several illustrations in which two collections of data have the **same mean** but **different variability**. It is also possible for two collections of data to have **different means** and the **same variability**. For example, consider the two graphs shown below. Notice that the mean of the data in Graph ________ is smaller than the
A / B

A

mean of the other collection of data.



GRAPH A GRAPH B

39. However, the variability of the two collections of data (shown above) about their respective means ________
is/is not

is

almost identical. Thus, the two distributions could be described as having similar _________________ but
variability/means

variability

different ______________.
variability/means

means

239. You could look at this same characteristic of the sampling distribution from a different point of view. Consider the question: How many samples of size 2 would have led to an absolute error of estimation of **1 or less**? Only samples with means of 9, 10, or 11 would lead to an absolute error of estimation of 1 or less. In the 100 samples of size 2, there were 19 + 20 + 16 samples with means that would have led to an absolute error of estimation of ______ or less. 1

240. Similarly, considering the experimental sampling distribution of sample means from samples of size 5, there were ______ samples that would have led to an absolute error of estimation of 1 or less. 76

241. The answer to the previous frame was obtained by adding the frequencies of samples with means of 9, 10, or 11. These 3 frequencies were ______, ______, and ______. Therefore, the sum of these frequencies was 76. 23, 32 21

242. According to the experimental sampling distribution for samples of size 25, there were ______ samples that would have led to an absolute error of estimation of 1 or less. 99

243. The answer to the previous frame was obtained by adding the frequency for sample means of 9, 10, and 11 — in other words, by adding ______, ______, and ______. 15, 67, 17

244. Finally, when the sample size was increased to 100, there were ______ samples which would have led to an absolute estimation error of 1 or less. 100

40. Therefore, simply knowing that two collections of data are similar in terms of their central tendencies ____________ (does/does not) indicate whether they are similar in terms of their variability. — does not

41. Another illustration of the lack of any relationship between the mean of a distribution and its variability is given by the following two tables of raw data:

Observation	Value	Deviation From 4
1	6	2
2	4	0
3	2	-2

TABLE A

Observation	Value	Deviation From 8
1	10	2
2	8	0
3	6	-2

TABLE B

Notice that the data in each table consists of _____ observations of a ____________ (numerical/non-numerical) variable. — 3; numerical

42. The particular reference value from which the deviations are calculated in each of the previous tables is the ____________ of that collection of data, since the sum of the deviations in each table equals ______. Thus, the mean of the data in Table A is _____, whereas the mean of the data in the Table B is _____. — mean; 0; 4; 8

236. (Continued)

Frequency distributions of 100 sample means based on samples drawn randomly from a large population. (The population has an approximately normal distribution, with a mean of 10 and a variance of 8.)

*Sample mean	Sample Size			
	2	5	25	100
5	1	0	0	0
6	2	0	0	0
7	8	1	0	0
8	11	10	1	0
9	19	23	15	3
10	20	32	67	93
11	16	21	17	4
12	13	9	0	0
13	6	3	0	0
14	3	1	0	0
15	1	0	0	0

* The mean of each sample was rounded off to the nearest value shown in column one.

237. Let us compare the four distributions in terms of the number of absolute errors of estimation greater than 1. For example, in the 100 samples of size 2, there were ______ samples that would have led to an absolute error of estimation greater than 1. 45

238. The answer to the previous frame was obtained by adding the frequencies 1, 2, 8, 13, 6, 3, and 1. These are the frequencies of all sample means of size 2 which would have led to an absolute error of estimation greater than ______, (i.e., which had sample means that differed from μ [10] by **more than** 1).

43. We could represent the **variability** of each collection of data by its range. The range of the data in Table A is _____ and the range of the data in Table B is _____. 4, 4

44. Earlier, we pointed out how the variability of a collection of data could be thought of as the degree to which the observed values were spread out or dispersed about the mean of that collection of data. The _________ range
is a useful measure of variability because it is the difference between the value having greatest **positive** deviation from the mean and the value having greatest **negative** deviation from the mean.

45. We also noted earlier that the range is **not** a completely satisfactory way of representing variability. Two distributions may have identical ranges and yet one distribution may appear to be much more variable than the other. For example, consider the two collections of data shown in the following tables:

Observation	Value
1	15
2	5
3	5
4	5
5	5

DATA # A

Observation	Value
1	15
2	5
3	12
4	7
5	8

DATA # B

The smallest value in each table is _______ and the 5
largest value is _____. Notice that the range of **both** 15
collections of data equals ______. 10

46. While the range is the same in both collections of data, all the values, except one, are identical in Data ______. A
A / B

232. Samples with means less than 7 or greater than 13 would have an absolute difference between $\bar{x}$ and μ greater than 3. Thus, there are only _______ samples of size 2 that would have lead to an absolute error of estimation **greater than 3.** 7

233. Only one sample of size 2 had a mean of 5. Two samples had a mean of 6, ______ samples had a mean of 14, 3
and one sample had a mean of _______. Thus, there 15
were, in all, ______ samples of size 2 that would have 7
led you to make an absolute error of estimation greater than 3.

234. Whereas there were 7 samples of size 2 which would have led to an absolute error of estimation greater than 3, only ______ sample(s) of size 5 would have led to an absolute error of estimation that large. 1

235. The one sample of size 5 you obtained that would have led to an absolute error of estimation greater than 3 was the sample with a mean of ______, since the difference 14
between this mean and the true mean was _____. 4

236. There were ______ samples of size 25 and of size 100 which would have led you to make an absolute error of estimation greater than 3. (We have repeated the previous table for your convenience.) 0 (no)

47. This illustration points up one of the disadvantages of the range as a way of representing variability. The range is determined by only two of the observed values in the collection of data:

1) the ______________ observed value, and — largest

2) the ______________ observed value. — smallest

48. Since the largest observed value in each table is ____ and the smallest observed value is _____, the collections of data shown in the following two tables have ________________ ranges. (the same/different) — 10 — 1 — the same

Observation	Value	Deviation From 5
1	1	-4
2	4	-1
3	5	0
4	5	0
5	5	0
6	10	4

TABLE A

Observation	Value	Deviation From 5
1	1	-4
2	1	-4
3	1	-4
4	8	3
5	9	4
6	10	5

TABLE B

49. We have listed the deviations of the values in each table from their common mean of ____. The largest and smallest values in both collections of data are the same. However, the other observed values in Table ____ (A/B) are all very close (or identical) to the mean. All the values in the other distribution are almost as far away from the mean as are the two extreme values. — 5 — A

230. (Continued)

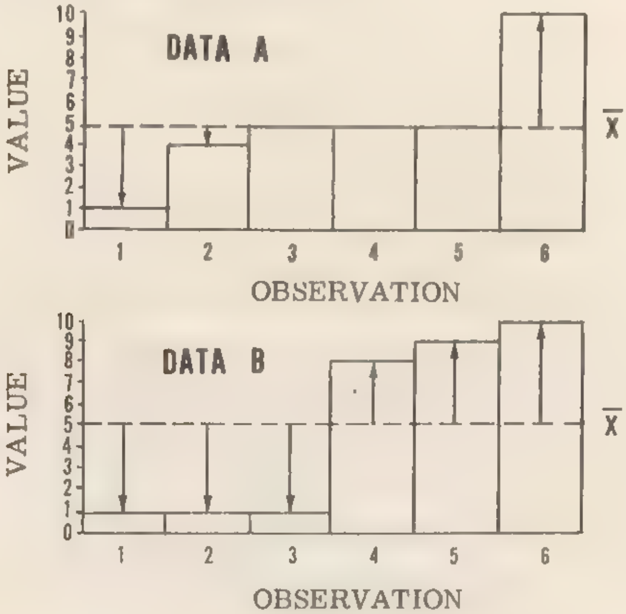
Frequency distributions of 100 sample means based on samples drawn randomly from a large population. (The population has an approximately normal distribution, with a mean of 10 and a variance of 8.)

*Sample mean	Sample Size			
	2	5	25	100
5	1	0	0	0
6	2	0	0	0
7	8	1	0	0
8	11	10	1	0
9	19	23	15	3
10	20	32	67	93
11	16	21	17	4
12	13	9	0	0
13	6	3	0	0
14	3	1	0	0
15	1	0	0	0

* The mean of each sample was rounded off to the nearest value shown in column one.

231. If you used each sample mean as an estimate of the population mean, you would have made an absolute error of estimation **greater than** 4 only two times in the 100 samples of size two. Only sample means of 5 or 15 would lead you to make an absolute error of estimation **greater than** 4. Only one sample of size 2 had a mean of 5 and only one sample of size 2 had a mean of 15. Thus, ________ samples of size 2 would lead to an error of estimation **greater than** 4. 2

50. Even though the two extreme deviations in both collections of data are the same, the typical size of a deviation in Data $\underset{A\,/\,B}{\underline{\qquad}}$ is greater than in the other collection of data. This feature can be illustrated by the following graphs of raw data: B



We have indicated the deviations from the mean in both collections of data with arrows, just as we have done previously. The lengths of these arrows indicate how most of the observed values in Data $\underset{A\,/\,B}{\underline{\qquad}}$ are very close to the mean, whereas the values in the other graph tend to be farther away from the mean. A

51. We have already seen that the mean of a collection of values can be thought of as the typical value. Therefore, one way we might represent the typical size of a deviation would be to find the mean, or the average of the deviations.

However, our formula for finding the mean of a collection of values ($\sum X/N$) says that the first thing to do is to add together all the values. No matter what the variability of a distribution, the sum of the deviations from the mean would always equal ____. 0

227. After rounding off (grouping) the sample means in this manner, we calculated the frequencies in each collection of 100 sample means. The preceding table lists these frequencies for samples of sizes ____, ____, ____ and ____. 2, 5, 25 100

228. According to the table, there are only two samples of size 2 with means (rounded off) of 5. There were also only two samples of size 2 that had means equal to 6. There were 8 samples of size 2 having means equal to ____, and 11 samples of size 2 with means equal to ___. 7, 8

229. The column on the far right of the table contains the experimental sampling distribution of sample means in samples of size ______. There were 93 samples of this size that had means of ______. 100 10

230. Notice that the most frequently occurring sample mean in each of the four experimental sampling distributions was a sample mean of ______. 10

You ______________ say, therefore, that the **modal** could
could/could not

sample mean in each of the four experimental sampling distributions was equal to μ. (The table we presented earlier is repeated here for your convenience. Refer back to this table whenever necessary as you answer the following frames.)

52. Statisticians have found it useful to represent the variability of a distribution by the typical or average size of the **squared** deviations from the mean. Whenever you square a number, your answer will be positive, whether the number you are squaring is positive or negative.

For example, $2^2 = 2 \cdot 2 = 4$

Similarly, $(-2)^2 = (-2)(-2) =$ _____ 4

Therefore, whether a deviation is positive or negative, when you square it, your answer will be _______________. positive
positive/negative

53. In the table of data shown below, we have listed 3 observed values, their deviations from the mean, and the square of these deviations. Thus, X_1 has a value of ____ and a deviation from the mean of ______. The 8, 3
square of that deviation is ______ times _______, which 3, 3
equals _______. 9

Observation	Value	Deviation from 5	Squared Deviation
1	8	3	9
2	-2	-3	9
3	5	0	0

If you add (sum) all 3 observed values, you obtain the **total** of the observed values, which is _____. Dividing 15
this total by ____ indicates that the mean of these three 3
values is ____. 5

54. Adding up the deviations from the mean value of 5 will naturally yield an answer of _____, because the mean is 0
defined as that particular reference value from which the deviations sum to _____. 0

224. (Continued)

Frequency distributions of 100 sample means based on samples drawn randomly from a large population. (The population has an approximately normal distribution, with a mean of 10 and a variance of 8.)

*Sample mean	Sample Size			
	2	5	25	100
5	1	0	0	0
6	2	0	0	0
7	8	1	0	0
8	11	10	1	0
9	19	23	15	3
10	20	32	67	93
11	16	21	17	4
12	13	9	0	0
13	6	3	0	0
14	3	1	0	0
15	1	0	0	0

* The mean of each sample was rounded-off to the nearest value shown in column one.

225. In the first column of the table, we have listed ______ possible values of the sample mean. The smallest value of the sample mean listed is ______ and the largest value listed is ______. 11 5 15

226. Each sample mean was rounded off to the nearest value appearing in the first column. In other words, if the sample mean were $6\frac{3}{4}$, we would consider it equal to a sample mean of 7. If a sample mean were $6\frac{1}{4}$, we would consider it equal to a sample mean of _____. 6

55. To find the mean of a group of values, you first find their total, and then divide this total by the number of values. If you wanted to find the mean (average) squared deviation, you could first add all the squared deviations and then divide by the number of squared deviations. In our previous example (see the table in Frame 53), the total of the number of squared deviations equals ______ plus ______ plus _____, which equals ____. 9, 9, 0, 18

Since there were 3 observations in the collection of data, you should divide the total of the squared deviations by _______ to find the mean of these squared deviations. 3

Thus, the mean (average) of the squared deviations equals _______ divided by _______, which yields the 18, 3
answer of ______. 6

56. Statisticians refer to the average of the squared deviations as the **variance**. Therefore, the **variance** of the previous collection of data is ______. 6

57. The **variance** is a statistic representing the variability of a collection of data. Values that are widely dispersed (spread out from the mean) have large deviations from the mean. Whether these deviations are positive or negative, they will result in large squared deviations. Therefore, saying that a collection has a large **variance** implies that the values tend to be ________________________ (spread out from/ close to) the mean. spread out fro

58. A collection of data having observations of all the same value has the ___________ (most/ least) possible variability. least

220. You could obtain random samples from the population of weekly mileages in the same manner in which you obtained random samples from the population of student opinions. Let's suppose you obtained 100 samples of size 2 in this manner. You could calculate the mean of each of these samples so that you would obtain ______ sample means. 100

221. We have already stated that we would represent the population mean by the symbol μ. We shall represent the sample mean by $\bar{x}$. Thus, if one of the samples of size 2 had a mean of 5, μ would equal ______ and $\bar{x}$ would equal ______, since we just said that the population mean was 10. 10 5

222. Having collected 100 samples, we would have 100 ______'s (μ / $\bar{x}$) but only one ______ (μ / $\bar{x}$). $\bar{x}$, μ

223. In a similar manner, we could obtain 100 samples of size 5 from the same population, calculate the mean of each sample, and the frequency distribution of these ______ sample means. 100

224. We could repeat this process to obtain 100 samples of size 25, and we could repeat it again to obtain 100 samples of size 100. The following table contains experimental sampling distributions similar to those you could expect to obtain for samples of size 2, size 5, size ______, and size ______. 25, 100

59. Suppose all the observations in a collection of data had the value 10. The mean of that collection of data would equal _____ and each observation would have a deviation from the mean equal to _____.

10
0

If each deviation equaled zero, each deviation squared would also equal zero. Since the **variance** of a collection of data is the typical size of the squared deviations, the variance of this collection of data would equal _________.

0

60. A collection of data having the least possible variability would have, therefore, a **variance** equal to _____. (It would also have a range equal to ____.)

0
0

61. Consider the collection of data shown below:

Observation	Value	Deviation From 6
1	8	
2	5	
3	5	
4	6	

The mean of this collection of data is 6. We have left room in the third column of the table to insert the deviation of each observation from the mean. The deviation of observation 1 from the mean equals _____ minus _____, which equals _____.

8, 6
2

62. Squaring the deviation of observation 1 from the mean equals ______ times _______, which equals _______.

2, 2, 4

218. In an earlier illustration, we considered 4 experimental sampling distributions, each of which was based on 100 samples obtained from a population of 10, 000 student opinions. The major difference in these 4 experimental sampling distributions was the size of the samples (the number of opinions in each sample). Let's consider another illustration of how the sampling distribution is altered by increasing the size of the sample.

Let's suppose you obtained 100 samples by a random sampling procedure from a large population. Let's suppose this population has an approximately normal distribution (the special type of bell-shaped distribution we discussed earlier), with a population mean equal to 10 and a population variance of 8. The symbol μ is often used to represent the **parametric mean**. Thus, in this illustration μ would equal ________. 10

Since we indicated earlier that we would use the symbol __________ to represent the population variance, $\tilde{\sigma}^2$ $\tilde{\sigma}^2$
$\tilde{\sigma}^2 / \sigma^2$
would equal _______ in this illustration. 8

219. Imagine that this population consists of the number of miles each of the 10, 000 students estimated he traveled to and from school each week. Lets suppose the students were asked to report this mileage **rounded off to the closest mile** on their registration forms. In other words you would have a weekly mileage from each of the 10, 000 students at the university. These 10, 000 mileages would be the __________________ from which we drew the samples. population

63. We could find the deviation and the square of that deviation in a similiar manner for each of the other observed values and summarize our work in the following table:

Observation	Value	Deviation From 6	Sq. of the Deviation
1	8	2	4
2	5	-1	1
3	5	-1	1
4	6	0	0

To find the average of the squared deviations (the variance), we would __________ all of the squared deviations and divide this total by _____. add, (sum) 4

Therefore, the variance of the previous collection of data equals _____ divided by ______. Thus, 1.5 is the _______________ of the previous collection of data. 6, 4 variance

64. The variance of the data in the following table equals _______ divided by _______. The variance, therefore, is ______. 100, 4 25

Observation	Value	Deviation from 10	Sq. of Deviation
1	15	5	25
2	5	-5	25
3	5	-5	25
4	15	5	25

65. Since the variance is the _______________ of the squared deviations from the mean, the **formula** for the variance will be similar in some ways to the formula for the mean we considered earlier. mean

66. To find the mean of a group of values, we first sum all the values and then divide this sum by the _____________ of values. number

213. (Continued)

heights as a sample of size ______ from the population of 10, 000 student ______________. The procedure used to obtain the sample would be a ____________________ sampling procedure. — 10 / heights / random

214. You could calculate any one of a number of statistics that would describe this sample of 10 heights. For example, you could find the ______________ of the sample by calculating the **difference** between the greatest height and the least height, in the sample, or you could add all the heights together and divide by 10. This would give you the ____________ of the sample, since there were 10 heights in the sample. — range / mean

215. Or you could use the following formula

$$\frac{\sum(X - \bar{x})^2}{N}$$

to calculate the _________________ of the sample. — variance

216. Suppose you obtained **100 samples of size 10** by a random sampling procedure from the previously described population. If you calculated the variance of each sample, you would have __________ sample variances. — 100

100/10

217. You could make a frequency distribution of these 100 sample variances. This distribution would be an experimental _______________ _________________ of sample variances based on 100 samples of size 10 obtained by a random procedure from this particular population. — sampling distribution

66. (Continued)

In a similar manner, to find the mean of a group of squared deviations, we first **sum** all the squared deviations and then **divide** this sum by the number of squared deviations. In other words, the total of all the squared deviations divided by the number of deviations equals the **mean** of the squared deviations, which particular mean we call the ___________. variance

67. Representing the mean of a group of values by $\bar{x}$ and the value of a particular observation by X, we would represent the **deviation** of that observation from the mean as (X – ___). The expression $(X - \bar{x})^2$ would represent the square of the previous deviation. $\bar{x}$

68. If your collection of data consisted of three observations, you could represent the _________ of these three observations as: sum (total)

$$X_1 + X_2 + X_3$$

69. Similarly, you could represent the sum of the _________ of these three observations from their mean as: deviations

$$(X_1 - \bar{x}) + (X_2 - \bar{x}) + (X_3 - \bar{x})$$

Finally, you could represent the sum of the ___________ deviations as: squared

$$(X_1 - \bar{x})^2 + (X_2 - \bar{x})^2 + (X_3 - \bar{x})^2.$$

70. Just as we represented the **sum** of all the raw scores by $\sum X$, we could represent the _________ of the deviations from the mean by $\sum(X - \bar{x})$. sum

211. Assuming that you felt the distributions of samples you obtained were representative of the distribution of samples you would obtain in the future, you might state your **confidence** more explicitly as follows:

> "I would expect to make an absolute error of estimation greater than .1 about 48 times out of 100 with a sample size of ______ drawn randomly from the same population, whereas I would only expect to make an absolute error that large about 32 times out of 100 if my sample size were ______."
> 2/10   2/10

2

10

212. A **confidence statement** of this sort is a very important type of statistical statement. While you cannot say for certain that increasing the sample size will increase the accuracy of a single sample, you know that, in general, your **confidence** in an estimate will be greater, the ______________ is the size of the sample.
larger/smaller

larger

213. It would be possible to obtain an **experimental** sampling distribution for any sample statistic. You would simply obtain as many samples as you wished by some particular sampling procedure, compute the value of a particular sample statistic for each sample, and then make a frequency distribution of these sample statistics. For example, suppose you had recorded the **height** of each of the 10,000 students at the university. You could write each student's height on a slip of paper, mix all of these slips together in a basket, and draw out 10 heights (10 slips of paper). You could view these 10

71. Regardless of what collection of data we are describing, however, we know that $\sum(X - \bar{x})$ will always equal _____, according to our definition of the mean. 0

72. Using the symbol $\sum$ (the capital Greek letter __________) in the same way as we did when we wrote $\sum X$ and $\sum(X - \bar{x})$, we could represent the **total** of all the squared deviations by ______ $(X - \bar{x})^2$. sigma $\sum$

73. Since the variance is the **mean** of the squared deviations, we could represent the ______________ of a collection of data as: variance

$$\frac{\sum(X - \bar{x})^2}{N}$$

74. Statisticians use the symbol σ^2 to represent the variance. Using this symbol, a formula for the variance could be written as follows:

$$\_\_\_\_\_\_\_\_ = \frac{\sum(X - \bar{x})^2}{N}$$

σ^2

The symbol σ is the uncapitalized form of the Greek letter sigma. The summation symbol _______ was the capital Greek letter sigma. For this reason, the variance, represented by σ^2, is often referred to as s____________ squared. $\sum$ sigma

206. (Continued)

On all four graphs we have shaded columns corresponding to frequencies of samples which would have led to an absolute error of estimation **greater than** ________ in all four graphs. .1

207. We have not shaded columns associated with sample proportions of .4, .5 and .6, all of which would lead to an error of estimation that was either 0, −.1, or +.1. Thus, the **unshaded** columns represent the frequency of samples leading to an **absolute error of estimation** of _____ or less. .1

208. A comparison of these four distributions shows clearly that the **risk** of an absolute error of estimation greater than .1 was smallest for samples of size ______ and largest for samples of size ______. 100 2

209. In the light of this illustration, it appears that the larger the size of the sample, the ________ (more/less) confident you can be that the sample proportion will be close to the parametric proportion—assuming, of course, that you accept the distribution of these 100 sample statistics as representative of what you would obtain if you continued sampling. more

210 On the basis of these four experimental sampling distributions, you would probably expect to make an error of estimation greater than .1 **more** often if the sample size were ________ (large/small) than if it were ________ (large/small). small large

75. We now have defined **two formulas**. The first formula is:

$$\bar{x} = \frac{\sum X}{N},$$

where $\bar{x}$ (called the ____________) is a statistic representing the c____________ t______________ of a distribution. — mean; central tendency

The second formula is:

$$\sigma^2 = \frac{\sum(X - \bar{x})^2}{N},$$

where σ^2 (called the v____________) is a statistic representing the ______________ of a distribution. — variance; variability

76. Let's try using the formula for both the mean and the variance on the collection of data shown in the following table.

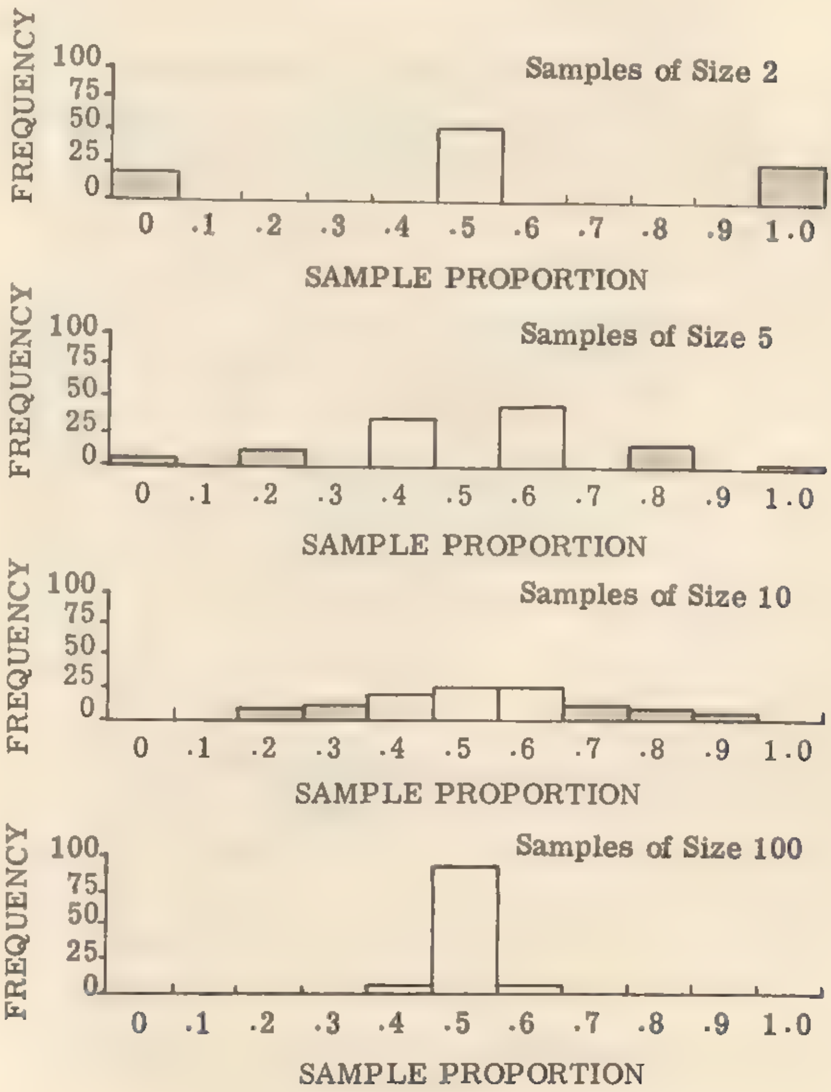
Observation	Value	$(X - \bar{x})$	$(X - \bar{x})^2$
1	4		
2	6		
3	5		

Notice that we have added two extra columns to a table of raw data. In the **third** column of the table, we could list the __________________ of each observation from the mean. In the **fourth** column we could list the ______________ of each deviation from the mean. — deviation; square

77. Before you can calculate the deviation of each value from the mean, you must first calculate the mean itself. Since the formula for the mean is __________, you must first find the __________ of all three values and then divide this result by ____________ (since N = 3). — $\frac{\sum X}{N}$; sum (total); 3

205. The remaining two graphs indicate the experimental sampling distribution for samples of size 10 and samples of size 100. It is obvious that the **variance of sample proportions** for samples of size 10 is ____________ (larger/ smaller) than the variance for samples of size 100, even though the **central tendency** of both distributions appears to be very close to the parametric proportion of ______. larger .5

206. A clear indication of how the **risk** of obtaining a highly **un**representative sample diminished as the size of the sample was increased is provided by the following figure.



78. The sum of the observed values is ________. Dividing this sum by 3, you find the mean to be ______. 15 5

79. Accordingly, you could replace the headings on columns 3 and 4 in the previous table as follows:

Observation	Value	(X - 5)	$(X - 5)^2$
1	4		
2	6		
3	5		

Note that here we have replaced ______ found in the earlier table with its actual value, ______. $\bar{x}$ 5

Notice that since the mean is 5, the deviation of observation 1 from the mean equals _______, and the square of this deviation equals ______. -1 1

The deviation of observation 2 equals _____, and the square of this deviation equals _____. 1 1

Finally, the deviation of observation 3 equals _______, and the square of this deviation equals _______. 0 0

80. We have summarized these answers in the following table:

Observation	Value	$(X - \bar{x})$	$(X - \bar{x})^2$
1	4	-1	1
2	6	+1	1
3	5	0	0

202. (Continued)

Graph A shows the distribution of sample proportions for samples of size 2. The bottom of the graph indicates the 11 sample proportions indicated in the first column of the previous table. Thus, the first column of the Graph A indicates the frequency with which samples having a sample proportion of 0 occurred in the 100 samples of size 2. The height of the column above the sample proportion of .5 in Graph A indicates the ____________ of samples with sample statistics of _____ in the 100 samples of size 2.

frequency
.5

203. Whereas only ______ of the 11 proportions listed along the bottom of Graph A are **possible** values of the sample proportion for samples of size 2, we have listed all 11 proportions in order to make it easy to compare the 4 experimental sampling distributions.

3

204. The distribution in Graph B is that of samples of size _______. Once again, we listed all 11 values of the sample proportion, even though there are only _____ possible values for samples of this size.

5
6

The most frequently occurring sample of size 5 is indicated by the highest column in the graph. In other words, the most frequently occurring sample proportion for the samples of size 5 was a proportion of ______, which occurred _____ times.

.6
40

80. (Continued)

Remember, the formula for the variance,

$$\sigma^2 = \frac{\sum(X - \bar{x})^2}{N},$$

says to sum all of the squared deviations and then to ____________ this sum by the number of observations. divide

The variance of the data in the previous table, therefore, equals ______ divided by ______. This means $\sigma^2 =$ ____ for this collection of three observed values. 2, 3, $\frac{2}{3}$

81. Suppose the deviation from the mean of every value in a collection of data equaled 2 or -2. The square of each of the positive deviations would equal 2 times 2, or 4, and the square of every negative deviation would equal _______ times ______, which would also equal 4. -2, -2

Since the square of every deviation from the mean would be 4, the typical or mean deviation would equal 4. We would, therefore, say that the **variance** of the distribution was _______. 4

82. Find the error in the following table.

Observation	Value	$(X - \bar{x})$	$(X - \bar{x})^2$
1	10	-5	25
2	4	-1	1
3	1	-4	16

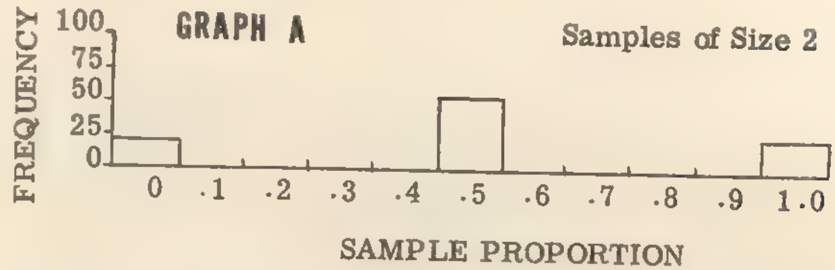
Notice that you would correct this error by changing _____ to ______, since the deviation of 10 from the reference value 5 is _________. -5, +5 (or simply 5) +5

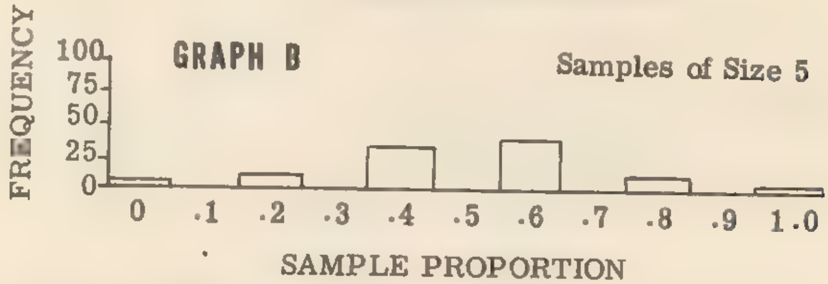
+5 / -5

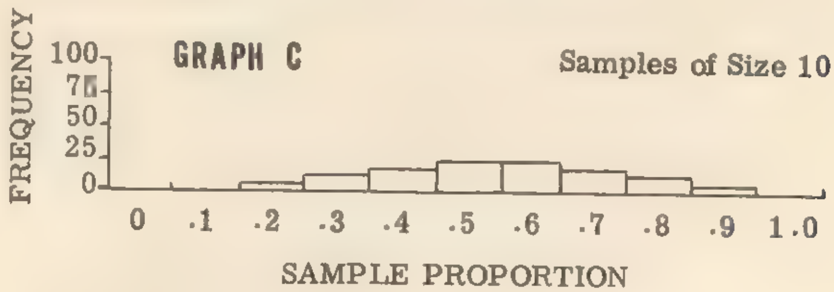
201. You could summarize the differences between the 4 experimental sampling distributions as follows:

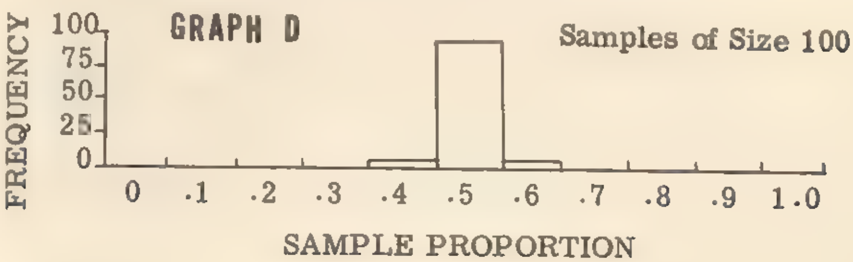
The larger the size of the sample, the ___________ risk there appears to be (greater/less) of making a large error of estimation. less

202. The differences between these 4 experimental sampling distributions are made even more apparent if we consider the **graphs** of these distributions.









83. Since the sum of the squared deviations equals ______ 42
and since there are 3 observations, the variance of
the previous data would equal _____ devided by _____. 42, 3
Therefore, σ^2 = _______. 14

84. If your data consisted of the values 3 and 7, the mean
would be _________. The variance, therefore, would 5
equal ______ divided by _______. In other words, 8, 2
σ^2 = _______. 4

85. Since $\bar{x} = \frac{\sum X}{N}$, the mean of 3 and 7 would equal $\frac{3 + 7}{2}$
or 5.

Similiarly, if the mean equals 5, $\frac{\sum(X - \bar{x})^2}{N}$ would equal
$\frac{(3 - 5)^2 + (7 - 5)^2}{2}$, which equals
$\frac{(-2)^3 + (2)^2}{2} = \frac{8}{2}$ or ______. 4

86. Since the variance can be thought of as the typical size of a **squared** deviation, statisticians have found it useful to assign a special name to the **square root** of the variance. They call the square root of the variance the **standard deviation**. Therefore, the standard deviation of a distribution is a deviation which, if squared, would
equal the _________________ (even if none of the observed variance
values actually has this particular deviation).

87. If the variance of your data were 9, an observation which
had a positive deviation from the mean of _____ would 3
have a squared deviation equal to the variance.

197. There were ______ samples of size 5 that would lead to an absolute estimation error greater than .2. (Remember, only values of p equal to **more than** .7 or **less than** .3 differ from the population proportion of .5 by more than .2.) 24

198. The 24 samples of size 5 that would had lead to an absolute error of estimation greater than .2 are made up of the two samples having a sample proportion of 0, the 10 samples having a sample proportion of .2, the 11 samples having a sample proportion of .8, and the ______ 1
sample(s) having had a sample proportion of ______. 1

199. While there were 24 samples of size 5 that would lead to an absolute error of estimation greater than .2, there were only ______ samples of size 10 that 10
would have led to an error so large.

The only samples of size 10 in which the sample proportion differed from the parametric proportion by more than .2 were the one sample in which p equaled .9, the 4 samples in which p equaled .8, and the ______ 5
samples in which p equaled ______. .2

200. Even more accurate estimates are apparently obtained from samples of size 100, since there were ______ 0
samples of size 100 in which the sample proportion differed from the population proportion by more than .2.

In fact, there were only ______ samples of size 100 5
that did not have a sample proportion identical to the population proportion. (Remember, the sample proportions were rounded off to the nearest $\frac{1}{10}$.)

88. In other words, if the variance of your data were 9, the **standard deviation** would be ______, since $(3)^2 = 9$. 3

If the variance of your data were 25, the standard deviation would equal _____, since $(\_\_\_\_)^2 = 25$. 5, 5

89. Just as a group of values may have a mean that is not equal to any of the values, it is not necessary for any particular value in a collection of data to have a deviation from the mean exactly equal to a standard deviation. Therefore, whereas the standard deviation can be thought of as the typical size of a deviation in same collection of data, it __________ (is/is not) necessary for any value in the data to actually have this deviation. is not

90. To summarize, the **variance** represented by the symbol ____________ is the typical or average squared deviation. σ^2

Any squared deviation equal to this average squared deviation would be called the ______________ ___________ standard deviation

Remember, it ____________ (is/is not) necessary that any value in the data actually have this particular deviation. is not

91. The **variance** is equal to the **square** of the **standard deviation**. Therefore, the standard deviation is equal to the ____________ ____________ of the variance. square root

Since the variance is represented by σ^2, the standard deviation could be represented by σ^2. But σ^2 is simply σ. Thus, the **variance** is represented by _______ (σ / σ^2) σ^2

and the **standard deviation** by ________ (σ / σ^2). σ

194. (Continued)

perfect estimate of the population parameter, most estimates would be pretty close to the population value. While you would have made 48 absolute estimation errors greater than .4 with the samples of size 2, there were only _____ samples of size 5 which would have led to an estimation error that large. 3

195. Among the 100 samples of size 5, there are only two samples with sample proportions of 0 and one sample with sample proportions of 1. Increasing the sample size to 5, therefore, appeared to ______________ the risk of making an absolute error of estimate greater than .4. reduce

reduce/ increase

196. (Once again, we repeat the table shown earlier for your convenience.)

TABLE A

Distribution of 100 samples drawn at random from a large population where $\tilde{p}$ equals .5.

Sample Proportion	Sample Size			
	2	5	10	*100
0	23	2	0	0
.1	-	-	0	0
.2	-	10	5	0
.3	-	-	12	0
.4	-	36	21	2
.5	52	-	25	95
.6	-	40	24	3
.7	-	-	10	0
.8	-	11	4	0
.9	-	-	1	0
1.0	25	1	0	0

* Sample proportions rounded off to the nearest one-tenth.

92. By using σ^2 to represent the __________ and σ to represent the __________ __________, we emphasize the fact that the variance is simply the __________ of the standard deviation. In other words, we emphasize that the __________ __________ of the variance equals the standard deviation by letting ______ represent the standard deviation and ______ represent the variance.

variance
standard deviation
square
square root
σ, σ^2

93. Suppose all the values in a collection of data deviated from the mean by either +3 or -3. The **variance** of that collection of data would equal ______ and the standard deviation would equal ______.

9
3

In other words, σ^2 equals ______ and σ equals ______.
9/3 9/3

9, 3

94. We have considered three statistics used to represent the central tendency of a collection of data. These three statistics are the __________, the __________, and the __________.

mean, median
mode

95. The __________ is the most frequently occurring value in a collection of data. The __________ is a value which would divide a list of the ranked data into two equal parts. The __________ is that particular reference value from which the sum of the deviations equals 0.

mode
median
mean

96. We have also considered three statistics used to represent the variability or dispersion of a collection of data. They are the __________, the __________ and its square root, called the __________ __________.

range, variance
standard deviation

191. Considering only the distribution based on samples of size 2, notice that _______ of the 100 samples had a sample proportion identical with the parametric proportion of .5. 52

192. If you felt the 100 samples of size 2 obtained were representative of the type of samples you would obtain through repeated sampling, you might make the following statement about samples of size 2 drawn from a population having a value of $\tilde{p}$ equal to 5: If you were estimating $\tilde{p}$ on the basis of a sample of size 2 from a population where $\tilde{p}$ was equal to .5, you would expect to make a perfect estimate about ______ times out of every 100 samples. 52

193. On the other hand, you could also make this following statement: You could expect to make an absolute estimation error **greater than** .4 about ______ times out of 100. 48

You would make an absolute estimation error greater than .4 if you obtained a sample proportion of either 0 or 1. Since there were 23 samples with sample proportions of 0 and _____ samples with sample proportions of 1 in the 100 samples of size 2 obtained, you would expect to make an estimation error that large about 48 times out of 100 (assuming, of course, you felt the previous 100 samples were representative of those you would expect to obtain in the future). 25

194. When you increased the sample size to 5, it was impossible to obtain a sample proportion of exactly .5. Thus, the frequency of sample proportions of .5 among samples of size 5 was necessarily ______. On the other hand, while samples of size 5 would never yield a 0

97. The ______________ is the difference between the largest and smallest observed value. The ___________ is the typical or average of the squared deviations from the mean of that collection of data. The standard deviation is a deviation which when squared will equal the ___________________.

range

variance

variance

98. We have also found a way to write rules for calculating the mean and the variance. To find the typical value or mean, we ____________ all the values and ____________ by the number of values. The formula for the mean is written, therefore, as:

add (sum), divide

$$\overline{X} = \_\_\_\_\_\_\_\_\_\_$$

$$\frac{\sum X}{N}$$

99. To find the typical squared deviation (the variance), we sum all of the ____________ _________________ and divide by the number of values. The formula for the variance, therefore, is written as:

squared deviations

$$\sigma^2 = \_\_\_\_\_\_\_\_\_\_$$

$$\frac{\sum(X - \overline{x})^2}{N}$$

The **standard deviation** is simply the __________ _______ of this variance.

square root

100. The statistics which describe the _________________ _________________ of the distribution represent in one way or another the typical value to be found in that distribution. A statistic representing the _____________ of the distribution describes the degree to which the observed values are spread out or dispersed.

central tendency

variability

188. The most frequently occurring sample proportion for samples of size 2 was ______, and the most frequently occurring sample proportion for samples of size 10 was _______. .5 .5

189. For your convenience, the following table is a duplicate of the one shown earlier.

TABLE A

Distribution of 100 samples drawn at random from a large population where $\tilde{p}$ equals .5.

Sample Proportion	Sample Size 2	5	10	*100
0	23	2	0	0
.1	-	-	0	0
.2	-	10	5	0
.3	-	-	12	0
.4	-	36	21	2
.5	52	-	25	95
.6	-	40	24	3
.7	-	-	10	0
.8	-	11	4	0
.9	-	-	1	0
1.0	25	1	0	0

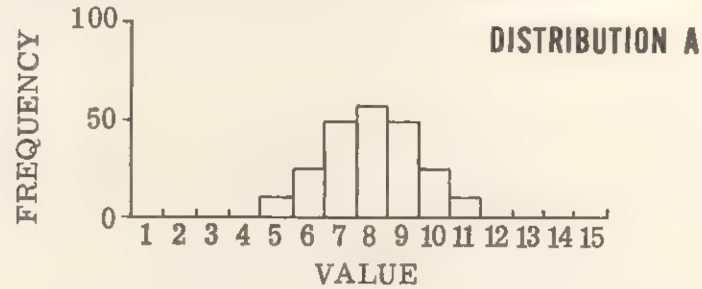
* Sample proportions rounded off to the nearest one-tenth.

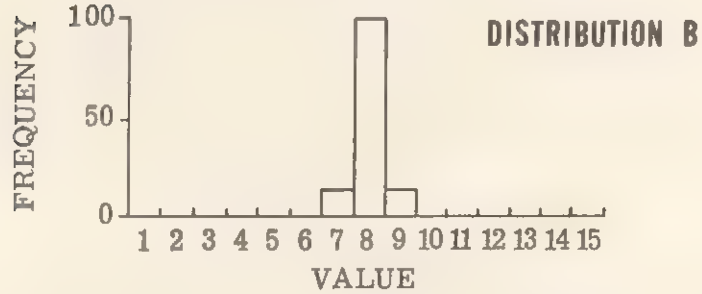
In terms of central tendency, in fact, all four sampling distributions seem to be clustered around the parametric proportion of _______. .5

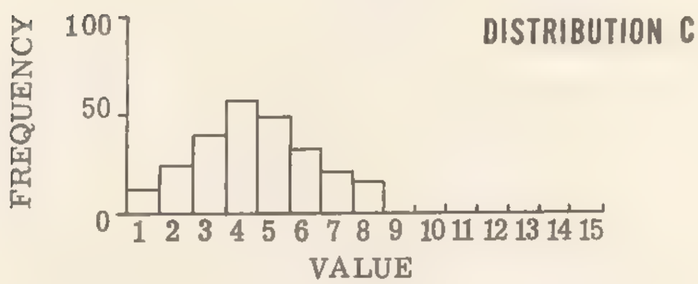
190. An important feature of these four distributions is the fact that the sample proportions were **less** variable in samples of size ________ then they were in the samples 100
2/100

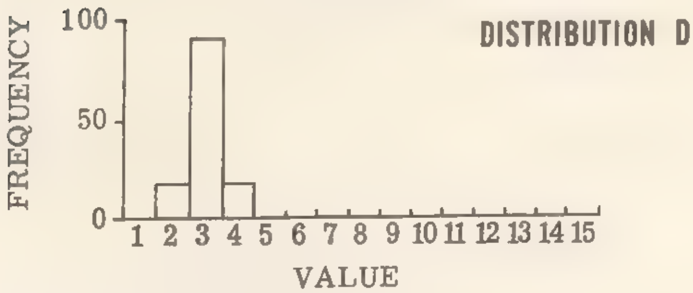
of size ________. 2
2/100

101. Statistics describing central tendency do **not** tell you anything about the variability of a distribution. Statistics describing variability do **not** tell you anything about central tendency. For example, consider the four distributions shown in the following graphs.









Notice that distributions A and B are similar in terms of their ____________________ (central tendency/ variability), yet quite different in terms of their ____________________ (central tendency/ variability).

central tendency

variability

184. If 34 of the 100 people had a favorable opinion, the actual sample proportion would be .34. This sample proportion would be between (less than one and greater than the other) the proportion .3 and the proportion ____ (.2/.4) that are shown in the first column of the table. .4

Since we rounded off the true value of p to the nearest $\frac{1}{10}$ when we calculated the frequency shown in the table, we considered a sample proportion of 34/100 to be a proportion of .3 (because 34/100 is closer to ____ (.3/.4) .3

than it is to ____ (.3/.4)). .4

185. A problem arises, however, when a sample proportion of 35/100 is obtained, since 35/100 is just as close to .3 as it is to .4. We simply made an arbitrary rule that if the sample proportion were exactly halfway between two of the proportions shown in the first column of the table, we would group it with the **smaller** proportion. In other words, we would consider a sample proportion of 35/100 to be equal to a proportion of .3 and a sample of 45/100 to be equal to ____ (.4/.5). .4

186. All eleven proportions shown in the first column were **possible** values of the sample proportion for samples of 100. According to the previous table, however, only three **observed** values of the sample proportion (rounded off to the nearest one-tenth) were .4, .5, and ____. .6

187. There were two samples of size 100 that had a sample proportion of .4 — 95 samples of this size having a sample proportion of ____ and ____ samples having a sample proportion of .6. .5, 3

102. On the other hand, Distribution A is similar to Distribution C in terms of ________, whereas the mean of Distribution A appears to be ________ the mean of Distribution C. variability

central tendency/variability

larger than/equal to

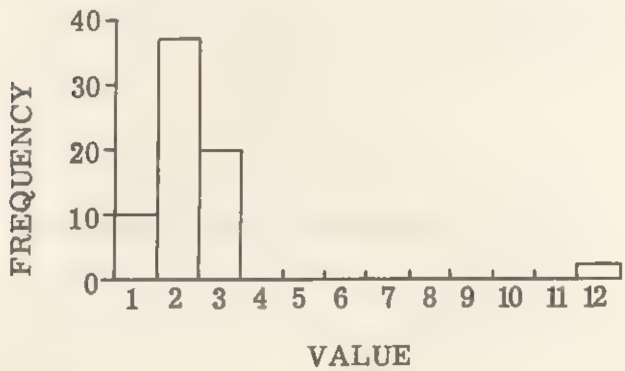
larger than

103. Distribution B and D are similar in terms of ________. variability

central tendency/variability

104. The choice of one statistic over another to represent same feature of a distribution depends upon the particular feature you wish to represent. For example, consider the distribution shown below. Except for a few extreme (unusually large) values, most of the values are grouped around the value ______. 2

2/5



If you represented the central tendency of the preceding data by the mean, the extreme values would tend to pull the mean away from the value 2. On the other hand, if you used the **mode** to represent the central tendency of the distribution, its value would be ______. In other words, the mode **might** be a better way of representing the central tendency in terms of this unusual distribution since the mode ________ influenced by the few extreme values found in the distribution. 2 / is not

is / is not

180. There are no dashes in the column containing the distribution based on samples of size 10. This is because every sample proportion listed in the first column of the table is a p__________ value of the sample proportion for samples of size 10. possible

181. The number in the same row as the proportion .2 and in the column headed by sample size 10 is _____. This indicates that exactly _______ samples of size 10 had a sample proportion of .2. 5 5

182. The column on the far right of the table is headed by a sample size of _______. 100

In a sample of 100, the smallest number of favorable opinions would be 0 and the largest number would be 100. Since there is a particular sample proportion corresponding to each possible number of favorable opinions, there are exactly ______ possible values of the sample proportion with a sample of 100 opinions. 101

183. In order to compare the sample distribution for size 100 with the sample distributions based on smaller samples, we have **grouped** the sample proportions in the samples of size 100 by **rounding off** each of sample proportions to the nearest $\frac{1}{10}$.

In other words, while a sample proportion of .02 would occur if two of the 100 opinions were favorable, we would consider that sample proportion to be 0, since .02 is closer to ______ than it is to _______. 0, .1
0/ .1 0/ .1

REVIEW II

FILL IN THE BLANKS:

1. The most frequently occurring value in the distribution is called the ______________. mode

2.

STUDENT	NUMBER OF ERRORS
A	3
B	6
C	2
D	3
E	11
F	8
G	3

In the table above, pick out the modal_____ number of errors. 3

3. A value in a collection of data which is smaller than half of the other observed values and larger than the remaining values is called the ____________. median

4. Find the median in the following collection of data:

   14, 11, 8, 4, 2 ______ 8

5. Find the median in the following set of numbers:

   9, 8, 6, 4 ______ 7

174. Therefore, you could say that of the ______ samples of size 2, there were _____ which had a sample proportion of 0, ______ which had a sample proportion of .5, and _______ which had a sample proportion of 1. These frequencies define an experimental sampling distribution of p for 100 samples of size 2.

100, 23, 52, 25

175. In the next column (headed by a sample size of 5), there are only _____ dashes. This indicates that of the eleven proportions listed in the first column only _____ could occur in samples of size 5.

5, 6

176. The frequencies 2, 10, 36, ____, ____, and ____ listed in the column for sample size 5 indicate the distribution of sample proportions for the 100 samples of size _____.

40, 11, 1. 5

177. By using the row headings to identify each value of the sample proportion, therefore, you could determine that exactly 2 of the 100 samples of size 5 had a sample proportion of 0 (no favorable opinions). Similarly, there were exactly 10 samples of size 5 that had a sample proportion of ______.

.2

178. There were 36 samples of size 5 that had a sample proportion of ____, and ____ samples of size 5 that had a sample proportion of .6.

.4, 40

179. The sum of the frequencies in the second column equals ________ because there were exactly 100 sample statistics in the distribution.

100

MULTIPLE CHOICE:

6. It is possible for two collections of data to have different means and:

   a. radial variability.

   b. diametric variability.

   c. the same variability.

   d. none of the above ______ c

7. The degree to which the observed values are spread out or are dispersed from the mean of a collection of data can be referred to as the:

   a. mean.

   b. variability of that collection of data.

   c. median.

   d. none of the above ______ b

8. The average of the squared deviations from the mean is called the:

   a. standard deviation.

   b. variance.

   c. mean.

   d. none of the above ______ b

9. The square root of the variance is called the:

   a. mean.

   b. median.

   c. standard deviation.

   d. none of the above ______ c

170. At the top of each of the 4 columns is a number indicating the size of the sample on which that particular distribution was based. The first of these headings is 2, indicating that the distribution shown in that column is based on samples of size 2. The second heading is _____, indicating that the distribution in that column is based on samples of size _____. 5 5

171. The last two columns indicate distributions of sample statistics based on samples of size _____ and samples of size _____, respectively. 10 100

172. The proportions listed in the column headed "Sample Proportion" represent different values of the sample statistic. For example, the second column contains the distribution based on samples of size 2. If you recall, there are only ____ possible values of the sample proportion if the sample size is 2. We have indicated this by placing a dash in each row of the second column corresponding to a sample proportion that _____________ (could/could not) occur with a sample of size 2. 3 could not

173. The **frequencies** 23, 52, and 25 appearing in the second column identify the distribution of sample statistics based on samples of size 2. There were 23 samples of size 2 in which there were **no** favorable opinions. This is why 23 appears in the same **row** as the sample proportion _____ (shown in the first column). 0

There were 52 samples of size 2 in which the sample proportion was _____, since 52 occurs in the same row as a sample proportion of .5. .5

Finally, there were _____ samples of size 2 in which all of the opinions were favorable (p equal to 1). 25

10. If the variance of your data were 36, the standard deviation would be:

   a. 9.

   b. 72.

   c. 4.

   d. none of the above ________ d

11. The symbol representing the standard deviation is:

   a. $\sum$.

   b. σ.

   c. σ^2.

   d. none of the above ________ b

TRUE OR FALSE:

12.

a negative deviation from the reference value.

________ false

13. Let us assume that the reference value is 8 and the observation is 6. In this case, the deviation is 2.

________ false

14. If we add all the deviations of a group of observed values from a particular reference value, and the answer is zero, that reference value is the mean of those values.

________ true

15. The difference between the largest and smallest value in a collection of data is called the median. ________ false

168. The following table presents four **experimental sampling distributions** similiar to those that would have been obtained using the procedure just described.

TABLE A

Distribution of 100 samples drawn at random from a large population where $\tilde{p}$ equals .5

Sample Proportion	Sample Size			
	2	5	10	*100
0	23	2	0	0
.1	-	-	0	0
.2	-	10	5	0
.3	-	-	12	0
.4	-	36	21	2
.5	52	-	25	95
.6	-	40	24	3
.7	-	-	10	0
.8	-	11	4	0
.9	-	-	1	0
1.0	25	1	0	0

* Sample proportions rounded off to the nearest one-tenth.

In the first column of the table, we have listed 11 values of the sample proportion, starting with a sample proportion of _____ (which indicates that none of the opinions 0
were favorable) and ending with a sample proportion of _____ (which indicates that all of the opinions were 1
favorable).

169. The 4 remaining columns of the table contain the distributions (frequencies) based on each of the 4 different sample sizes. The words __________ ________ sample size
printed at the top of the table refer to these four possible sample sizes.

16. If your data consisted of the values 73, 22, 14, 91, and 11, the range would be 47. ________ false

17. It is possible for two collections of data to have the same mean, but different variability. ________ true

163. It would also be possible to obtain 100 samples consisting of 100 opinions each. To do this you would merely draw _______ (slips of paper) from the basket each time you wished to collect a sample, remixing these slips in the basket before you drew the next sample.

100

164. If there were exactly 10 favorable opinions in the sample of 100, the value of the sample proportion (p) would be _______.

.1

165. To summarize, then, we have considered a way of obtaining 4 experimental sampling distributions. Each distribution consists of the __________________ for each possible value of the sample proportion that occurred in 100 samples.

frequencies

166. The major difference in these 4 experimental sampling distributions is the number of opinions included in each sample. We obtained one distribution for samples of size 2, another for samples of size ______, another for samples of size ______, and a fourth distribution for samples of size ______.

5
10
100

167. While the sample size on which each of these four distributions is based is different, the total of the frequencies in each distribution equals _______, since each distribution is based on _______ samples.

100
100

1. In addition to describing the **central tendency** and **variability** of a distribution, it is often useful to say something else about the general shape of the distribution. For example, consider first the distribution shown in Figure A below. Now look at Figure A^1 In Figure A, we have shown how half of Figure A would look reflected in a mirror.

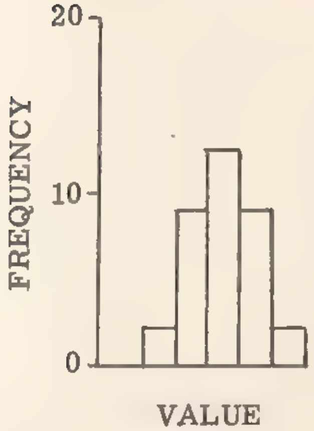


FIGURE A

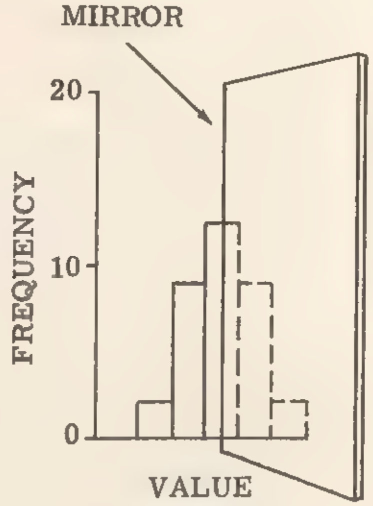


FIGURE A^1

Notice how the reflection of the **left** half of the distribution has exactly the same shape as the **right** half of the distribution. (The right half is behind the mirror in Figure A^1.) In other words, if we cut the distribution shown in Figure A in half, along the line where we placed the mirror in Figure A^1, we would have divided the distribution into two parts with the same shape but facing in opposite directions.

157. Notice that each sample of 5 could contain ____, ____, 0, 1
____, ____, ____, or ____ favorable opinions. 2, 3, 4, 5

158. Therefore, there are only _____ possible values of the 6
sample proportion: 0, .2, ____, ____, ____ and ____. .4, .6, .8, 1

159. If you determined the frequencies with which each of the possible values of the sample proportion occurred in the 100 samples, the total of all of the frequencies would
equal _______. 100

160. In the same manner as you obtained 100 samples of size 2 and 100 samples of size 5, it would be possible to obtain 100 samples of size 10. Instead of drawing 2 opinions for each sample or 5 opinions for each sample,
you would draw ______ opinions for each of the 100 10
samples.

161. The fewest possible favorable opinions in each sample of 10 would be 0. The largest number of favorable
opinions would be ______. 10

Therefore, there are _______ possible values of the 11
sample proportion: ____, ____, ____, ____, ____, ____ 0, .1, .2, .3, .4, .5
____, ____, ____, ____ and ____. .6, .7, .8, .9, 1

162. Samples of size 5 could not yield a sample proportion
of .1, whereas a sample of size _______ can have p 10
10/2
equal to .1.

1. (Continued)

In comparison, consider the distribution shown in Figure B below. In Figure B[1] the reflection in the mirror ____________ have the same shape as that
does/does not
part of the distribution behind the mirror.

does not

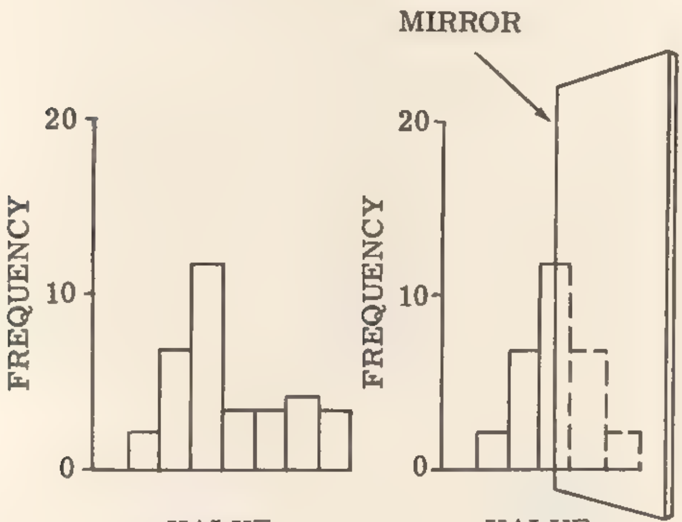


FIGURE B   FIGURE B'

Therefore, of the previous distributions A and B, Distribution ______ can be divided in half so that the
A/B
right half looks like a "**mirror image**" of the left half, whereas it would not be possible to divide the other distribution in such a manner.

A

152. Since the sample proportion (p) is simply the number of favorable opinions divided by the number of opinions in the sample, there are only 3 possible values of the sample proportion: _____, _____, and _____. 0, .5, 1

153. It ______________ be possible to obtain a sample
(would/would not)
proportion of .2 with a sample of size 2. would not

154. If you counted the number of samples (out of the 100 samples) that had a sample proportion of 0, a sample proportion of .5, or a sample proportion of 1, those frequencies would have to total _______. 100

155. The list of 3 frequencies indicating how often each of the 3 possible values of the sample proportion occurred in the 100 samples defines a __________________ of sample proportions. distribution

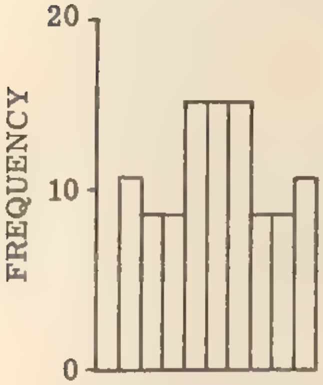
In other words, those 3 frequencies define a(n)
____________________ s________________ experimental sampling
(experimental/theoretical)
d________________ of 100 sample proportions. distribution

156. We now have an experimental sampling distribution of 100 sample statistics. Each sample statistic is the proportion of favorable opinions in a sample of size 2. It would be interesting to compare the experimental sampling distribution based on samples of size 2 with a similar sampling distribution based on larger samples. For example, you could obtain 100 additional samples in the same manner, this time including 5 opinions in each sample. These 100 samples of size _____ would provide you with _______ new sample proportions. 5 100

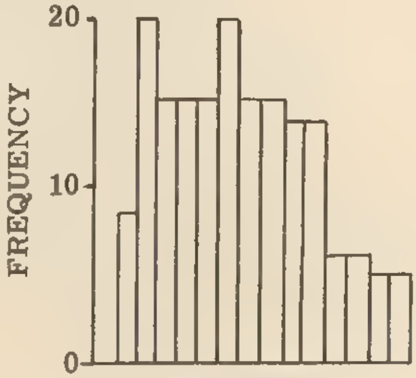
2. A distribution which can be divided so that its left half appears to be a mirror image of its right half is called a **symmetrical** distribution. If a distribution cannot be so divided, it is called an **asymmetrical** distribution. Thus, of the two distributions shown below, Distribution ____ (A/B) is symmetrical, whereas the other distribution is ______________.

A

asymmetrical

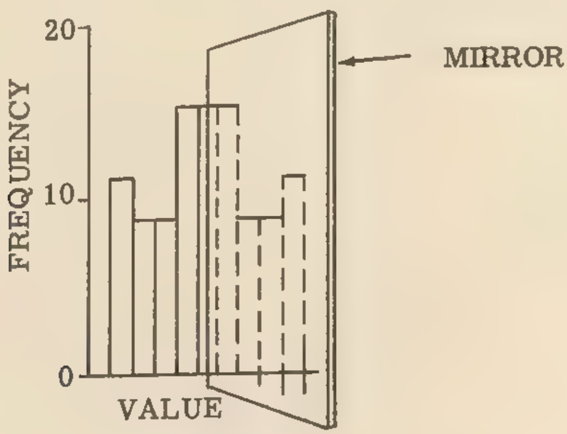


DISTRIBUTION A



DISTRIBUTION B

Distribution A would be called **symmetrical** since you could place a mirror so that the reflection in the mirror had exactly the same shape as that part of the distribution hidden behind the mirror. (See the illustration below.) In other words, the right half of the distribution is a "mirror image" of the left half of the distribution.



DISTRIBUTION A[1]

146. Since $\tilde{p}$ equals .5, we know that exactly ________ of the 10, 000 students have a favorable opinion concerning the overseas campus. 5, 000

147. There must be exactly 5, 000 students in favor of the overseas campus since $\tilde{p}$ equals .5 and 5, 000 divided by __________ equals .5. 10, 000

148. To begin with, imagine that 100 samples of 2 opinions each were drawn by the previously described **random** procedure from the population of 10, 000 student opinions. In other words, suppose the slips containing the 10, 000 opinions were mixed in the basket and exactly ______ slips were drawn out each time a sample was obtained. Further imagine that exactly ______ samples of two opinions each were obtained in this manner. (Assume the slips taken for one sample were remixed in the basket before the next sample was drawn.) 2 100

149. If we represented the proportion of favorable opinions in the first sample by p_1 and the proportion of favorable opinions in the second sample by p_2, the last sample proportion we obtained would be represented by ______. p_{100}

150. Each of these 100 sample proportions would have a particular value corresponding to the proportion of ______ responses in each of the 100 samples of size ___. favorable, 2

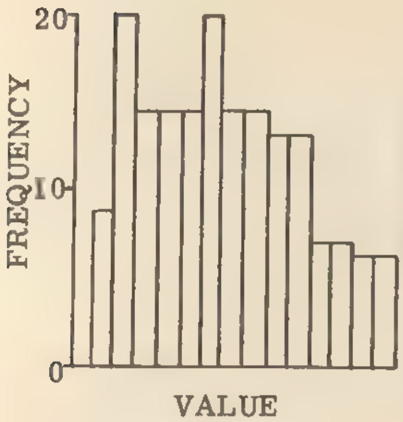
151. Notice that the **number** of favorable opinions in a sample of size 2 could only equal _____, _____, or _____. 0, 1, 2

2. (Continued)

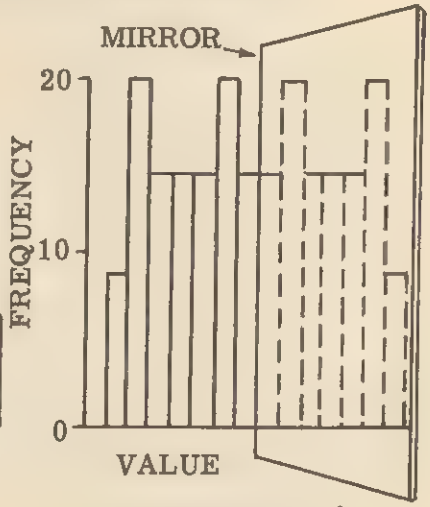
On the other hand, the reflection of the left half of Distribution B (Distribution B[1]) ____________ have the same shape as that part of the distribution hidden behind the mirror. We would say, therefore, that Distribution B was ______________________.

does/does not — does not

symmetrical/asymmetrical — asymmetrical



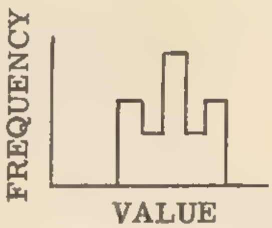
DISTRIBUTION B



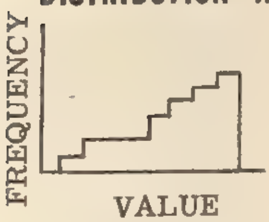
DISTRIBUTION B[1]

3. Of the four distributions shown below, Distributions _____ and ______ would be described as **symmetrical**, while the other two distributions would be described as **asymmetrical**. B, D

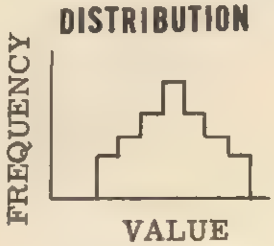
DISTRIBUTION A



DISTRIBUTION B



DISTRIBUTION C



DISTRIBUTION D

143. Sampling distributions based on an actual collection of samples (such as the ones we considered in the preceding illustrations) are often referred to as **experimental** sampling distributions. A sampling distribution based solely on logical considerations, therefore, would be called a(n) ______________ (experimental/theoretical) sampling distribution, whereas a sampling distribution based on a group of samples you had actually collected would be called a(n) ______________ (experimental/theoretical) sampling distribution.

theoretical

experimental

144. An important feature of sampling distributions is the relationship between the **variability** of the sampling distribution and the **size** of the samples. In general, the larger the size of the sample, the less the sample statistic will tend to vary from sample to sample. In other words, as you increase the size of the sample, you tend to ______________ (increase/decrease) the variance of the sampling distribution.

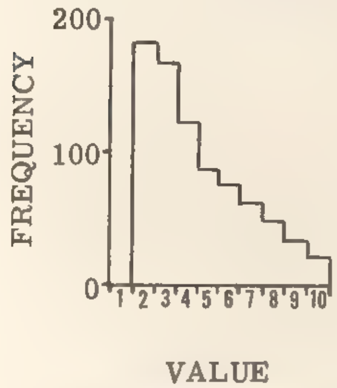
decrease

145. In order to illustrate the relationship between the size of the sample and the variability of the sampling distribution, let's consider the following example.

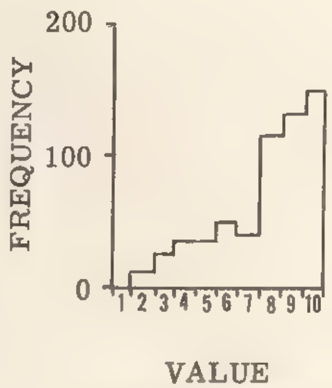
In the previous illustration, it was found that the proportion of the 10, 000 students who were in favor of the overseas campus was exactly .7. For the present illustration, let's suppose that the true population proportion of students in favor of the overseas campus was exactly .5. Thus, $\tilde{p}$ now equals _____ instead of .7, as was the case in the previous example.

.5

4. In some distributions, the scores seem to be piled up towards one end of the distribution. For example, in Distribution $\underset{\text{A/ B}}{\underline{\qquad}}$ (shown below) most of the observed values occurred near the **low-valued** end of the distribution, with only a few values occurring near the upper end of the distribution. A



DISTRIBUTION A



DISTRIBUTION B

5. The scores in Distribution B tend to be piled up near the $\underset{\text{high/ low}}{\underline{\qquad\qquad}}$ values, with only a few observed values down near the low end of the distribution. high

6. Both of the distributions shown above would be referred to as $\underset{\text{asymmetrical/ symmetrical}}{\underline{\qquad\qquad\qquad\qquad}}$ distributions. asymmetrical

7. Asymmetrical distributions in which most of the scores are piled up near one end could be regarded as "lopsided."

139. We already noted that 35 out of the 100 samples of size 5 would have led to an absolute error of estimation **greater than** .2. In order to make an absolute error of estimation greater than .2, you would have to obtain a sample of 10 that had **fewer** than ________ favorable 5
opinions or **more than** ________ favorable opinions. 9
(Remember, $\tilde{p}$ equals .7.)

140. While there were 35 samples of size 5 that would lead to an error of estimation greater than .2, there were only ______ such samples of size 10 (according to 5
the previous table).

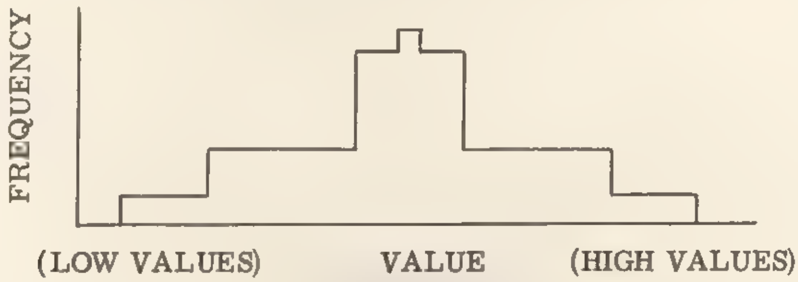
141. The only observed samples of size 10 in which the absolute difference between $\tilde{p}$ and p was greater than .2 were the 5 samples in which p equaled _______. .4

142. The sampling distribution for samples of size 5 and samples of size 10 were obtained in the illustrations by actually collecting these samples from a particular population by a particular procedure. Later, you will see how it is possible to **predict**, through logical considerations, the sampling distribution of a sample statistic obtained by a **random sampling procedure**. These **predicted** sampling distributions are often referred to as **theoretical** sampling distributions because they are based on logical considerations rather than on samples that have actually been collected.

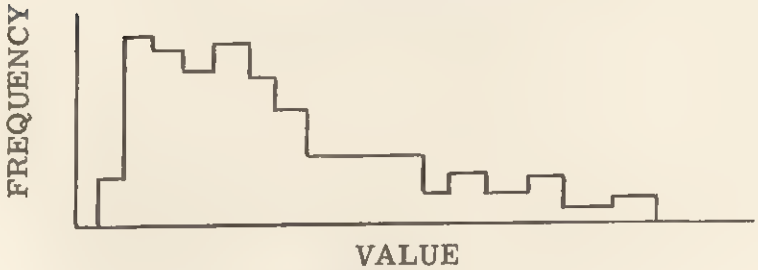
Thus, a sampling distribution based on an actual collection of samples _______________ be an example of would not
would/would not
a theoretical sampling distribution.

7. (Continued)

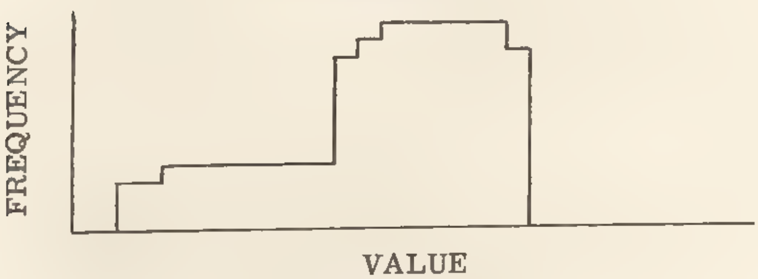
Distributions of this sort are said to be **skewed**. If a skewed distribution has most of the observations piled up near its **low** values, the distribution is said to be positively skewed. If the distribution has most of the observations piled up near the **high** values, the distribution is said to be **negatively** skewed. Thus, of the three distributions shown below, Distribution ___ B
is **positively skewed**, whereas Distribution _____ is C
negatively skewed. Distribution _____, however, is A
neither positively nor negatively skewed, since this distribution is symmetrical.



DISTRIBUTION A



DISTRIBUTION B



DISTRIBUTION C

136. In the illustration we just considered, it was suggested that a larger sample size would lead to fewer unrepresentative samples. The following table summarizes the distribution of sample proportions obtained when 100 samples of size 10 were collected by the **random sampling procedure** described earlier. The only difference in the procedure for collecting samples of size 10 and the procedure for collecting samples of size 5 is that you draw _______, rather than 5, slips of paper from the basket for each sample. 10

Number of favorable opinions in a sample of 10	Proportion of favorable opinions in the sample	Frequency of such samples
0	0	0
1	.1	0
2	.2	0
3	.3	0
4	.4	5
5	.5	10
6	.6	20
7	.7	40
8	.8	25
9	.9	0
10	1.0	0

137. Notice that when samples of 10 were collected, there were _______ **possible** values of the sample proportion. 11

138. Each of these 11 **possible** sample proportions corresponds to a particular number of favorable opinions in a sample of 10. (This number is shown in the first column of the preceding table.) Therefore, if there were 6 favorable opinions in the sample of 10, the sample proportion would be _________. .6

8. Remember that a ______ distribution is regarded as "lopsided." A ______ distribution, however, could be regarded as perfectly balanced at its center, so balanced that its left half is a "mirror image" of its right half.

(skewed/symmetrical; skewed/symmetrical)

skewed

symmetrical

9. We can think of any collection of data as a record of the observed ______ of a variable.

values

10. Any collection of data can be described in terms of the various frequencies with which the different values occur in the data. This group of frequencies is referred to as the ______ of the data.

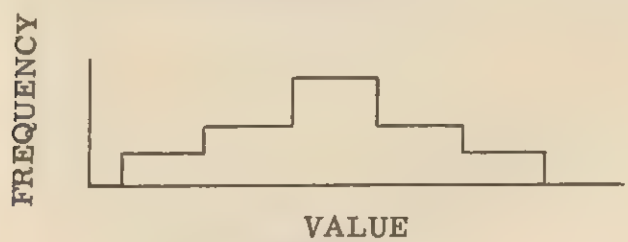
distribution

Often the difference between two distributions can be made apparent by drawing a picture of the distribution in the form of a frequency graph. Graphs of this sort make the "shape" of the distribution apparent.

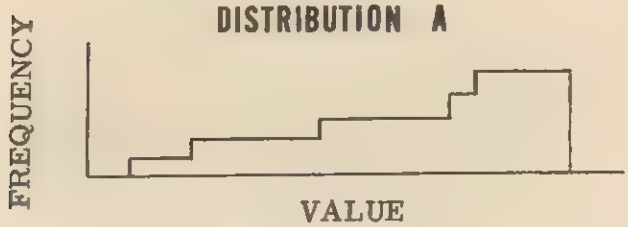
For example, we could describe Distribution ______ (A/B) (shown below) as a symmetrical distribution and Distribution ______ (A/B) as a ______ (positively/negatively) skewed distribution.

A

B, negatively



DISTRIBUTION A



DISTRIBUTION B

134. Knowing something about the distribution of sample statistics gives you some basis for deciding whether a sample is adequate for your purposes. According to the previous graph, only 35 of the 100 samples you obtained would have led to an absolute estimation error greater than .2. Or, to put it differently, only 35 sample statistics differed from the population statistic by more than ______. .2

If you believed that the distribution of sample statistics you had obtained was representative of the distribution you could expect in the future, you might make the following sort of statement:

> "If I obtained a sample of 5 opinions from this population by this random sampling procedure, I would expect to make an error of estimation greater than .2 only about ______ times out of 100." 35

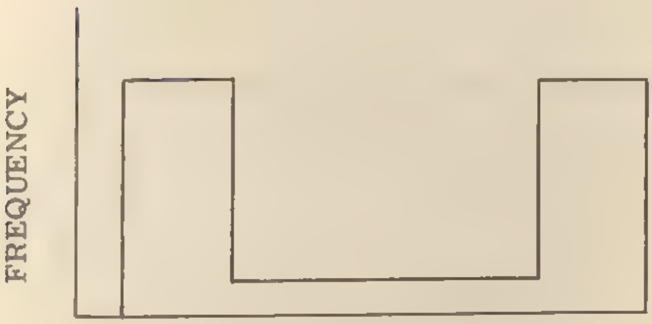
135. The previous statement is an example of one of the most important types of statistical reasoning. While you are not absolutely certain that your sample will give you a representative picture of the population, you can say something about the likelihood of making errors of estimation of any particular size. In other words, you can say something about your **confidence** in an estimate based on a single sample. A **confidence statement** of this sort is based upon your knowledge of the distribution of sample statistics obtained by a particular sampling procedure. In other words, it is based upon your knowledge of the s______ d______ of the sample statistics. sampling distribution

11. You would describe Distribution A (above) as **symmetrical** since its left half is simply the reverse of its right half. In other words, the left half of Distribution A mirrors its right half. Distribution B, however, is **not** symmetrical. Distribution B is "lopsided," with most of the observed values piled up near one end. Lopsided or **skewed** distributions are described as positively skewed when the piling up occurs near the ________ (high/low) -valued end of the distribution, and negatively skewed when the piling up occurs near the ________ (high/low) - valued end.

high

low

12.



Most of the observed values in the above graph are "piled up" near the ________ (ends/center) of the distribution, with very few observed values near the ________ (ends/center) of the distribution.

ends

center

13. The previous distribution, however, ________ (would/would not) be described as symmetrical, since its left side ________ (does/does not) mirror its right side.

would

does

131. (Continued)

Since there are only 5 opinions in each sample, there were only _______ possible values of the sample statistic: $\frac{0}{5}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}$, or $\frac{5}{5}$. — 6

132. We noted previously that most of the samples were reasonably representative of the population proportion $\tilde{p}$ equal to .7. (Samples with p equal to .6 or .8 were quite frequent.) There were, however, an appreciable number of samples in which p equaled .2, .4, or 1. Samples of this sort would give a(n) ______________ (representative/unrepresentative) impression of the population, since $\tilde{p}$ equaled ______. — unrepresentative; .7

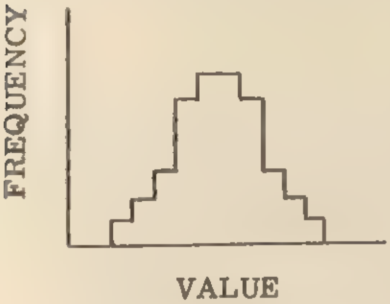
133. In order to decide whether to use a sample rather than to collect data on the whole population, it would be necessary to contrast the **advantages** of only obtaining a sample rather than a complete collection of data against the **risk** of obtaining an unrepresentative sample, one that would give you a distorted picture of the population. If it were very difficult to obtain a complete collection of data on the whole population and if you only needed a **rough estimate** of the population statistic, it might be adequate to base your estimate on a ______________. — sample

On the other hand, if it were not too expensive or too difficult to obtain complete information about the population and if you needed very accurate information about the population statistic, you probably ____________ (would/would not) estimate $\tilde{p}$ from p. — would not

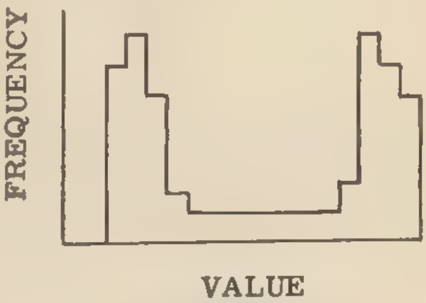
14. As a psychologist, you will encounter distributions with many different shapes. You will find, however, that certain types of distributions are encountered more often than others. For example, a very common type of distribution is one in which most of the values are piled up near the mean, with fewer and fewer values occurring farther from the mean.

In other words, values similar to the mean would have the ______ (larger/smaller) frequencies, whereas values farther away from the mean would have ______ (larger/smaller) frequencies. — larger; smaller

15. Distribution ____ (A/B) (below) would be an example of the common type of distribution we just described. — A

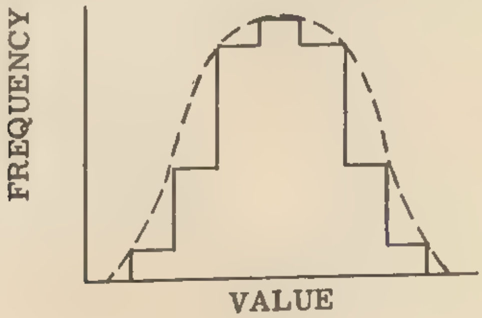


DISTRIBUTION A



DISTRIBUTION B

16. This commonly encountered type of distribution is often described as "bell-shaped" because its shape is similar to a bell's.



129. According to Table One, the proportion of favorable opinions in the second sample was ______. We could continue listing the proportion of favorable opinions in each successive sample (as indicated by the column of dots in this table) all the way to the 100th sample. The actual proportion of favorable opinions in the last sample, p_{100}, is equal to ______.

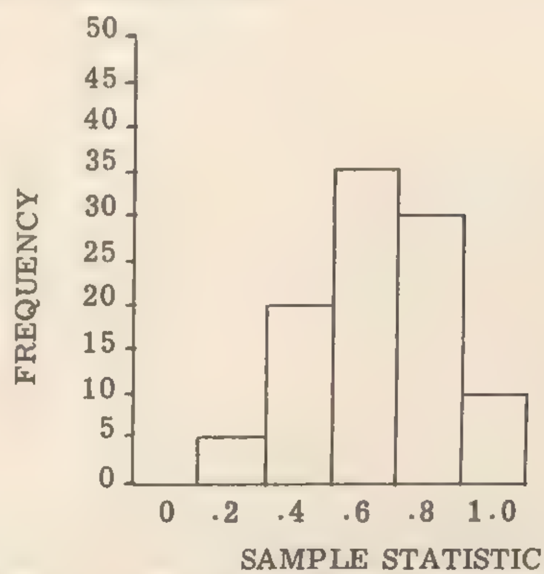
.5

.8

130. We are interested in a particular sample statistic: the **proportion** of **favorable** opinions in each sample. Sometimes this sample statistic had a large value, such as .8 or .9, and sometimes it had a low value, such as .5 or .2. Thus, the value of the sample statistic **varied** from one sample to the next. You saw how it was possible to calculate a **frequency distribution** for the observed values of the sample statistic (just as you would calculate a frequency distribution for the data in a table of raw data). The distribution of the sample statistic p is simply a list of the frequencies of each of the possible values of this s____________ s______________.

sample statistic

131. The following graph indicates the frequency distribution of the sample statistic p obtained from each of the 100 samples of 5 opinions each.

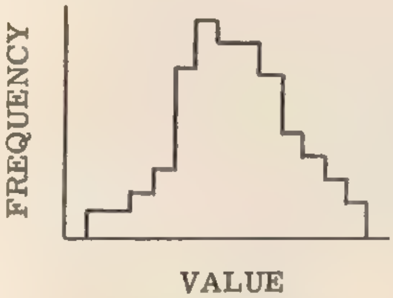


(PROPORTION OF FAVORABLE OPINIONS IN SAMPLES OF FIVE)

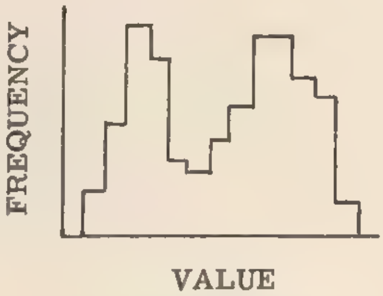
17. Although the shape of the previous distribution is **not** identical to the shape of a bell, it is convenient to describe this kind of distribution as **approximately** "bell-shaped." You could describe the difference between the two distributions below by saying that Distribution ______ (A/B) is approximately "bell-shaped," whereas Distribution ______ (A / B) is not.

A

B



DISTRIBUTION A



DISTRIBUTION B

18. Let's consider an example of a "bell-shaped" distribution. Suppose you filled a glass jar with 20 marbles and then asked a large number of students to estimate how many marbles there were in the jar. Some of the estimates would be too high and others would be too low. You would expect, however, that most of the estimates would be fairly close to the actual number of marbles in the jar. Occasionally, you would obtain some poor estimates, such as 15 or 25. On the other hand, you would expect estimates near ______________ (twenty/twenty-five) to be more frequent than estimates near ______________ (twenty/twenty-five).

twenty

twenty-five

126. Since we obtained 100 samples from the population, the proportion of favorable opinions in the last sample obtained could be represented by the symbol ________. p_{100}

127. Study the two tables shown below.

Table One

P	Value
p_1	.2
p_2	.5
p_3	.7
.	.
.	.
.	.
p_{99}	.3
p_{100}	.8

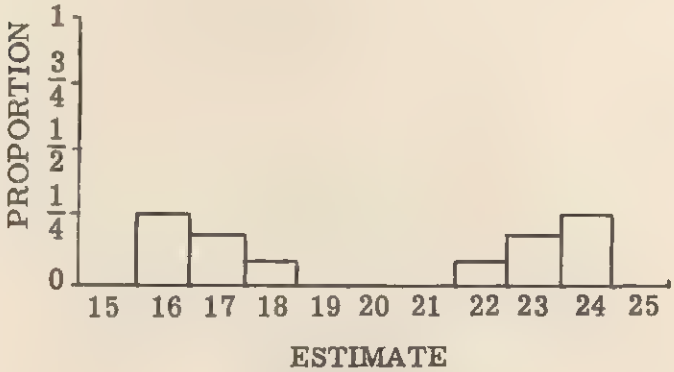
Table Two

Observation	Value
X_1	5
X_2	10
X_3	25
.	.
.	.
.	.
X_{99}	15
X_{100}	8

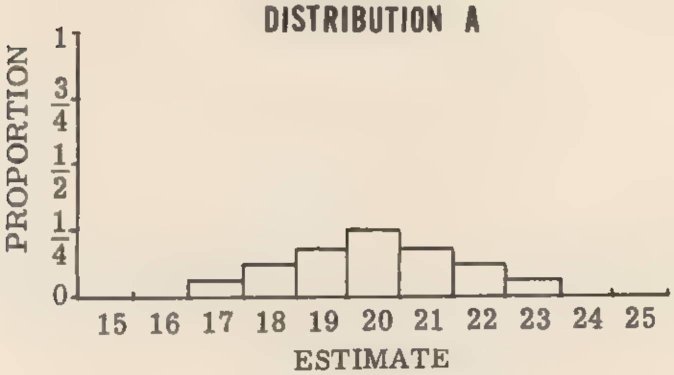
The table on the right (Table Two) is a table of r______ d__________ similar to the ones you saw earlier. We have indicated that the complete table contains ______ observations of a numerical variable. (We have simply used a series of 3 dots to indicate that the complete list would actually take up much more room.) raw data 100

128. The table on the left (Table One) is a very similar to a table of raw data. This is the sort of table in which we could list the proportion of favorable opinions in each of the 100 samples obtained. Thus, p_1 identifies the proportion of favorable opinions in the __________ sample. The actual p_1 was equal to .2 (as indicated by the entry in the first row and second column). The proportion of favorable opinions in the first sample was ______. first .2

19. You would probably find the distribution of these estimates similar to which of the distributions shown below? ______ B
A / B



DISTRIBUTION A



DISTRIBUTION B

20. Both of the above graphs would be called ______________ relative
absolute/ relative

frequency distributions, since they show the p__________ proportion
of subjects who estimated each of the possible values.

21. If Distribution A (above) had been the distribution of your data, the subjects would have been acting very strangely. According to Distribution A, many of the subjects over-estimated and many of the subjects under-estimated but very __________ of the subjects made estimates close to the true number of marbles in the jar. (Remember, there were 20 marbles in the jar.) few

22. According to the Distribution B, (above) **none** of the subjects' estimates was greater than ______________ 23
or less than ________. 17

121. The actual number of students in favor of the overseas branch in the population was .7. Thus, you could say that the population proportion of favorable opinions was ______. .7

122. In order to distinguish the proportion of favorable opinions in the **population** from the proportion of favorable opinions in a **sample**, we will represent the population proportion by the symbol $\tilde{p}$. Thus, in the present illustration, $\tilde{p}$ equals ____. .7

123. The little wavy line above the small letter p is called a "**tilde**" (pronounced til-da, so that it rhymes with Hilda). In general, whenever you see a tilde placed over a statistic, that statistic describes the **population**. You would know, therefore, that ______ represented the variance of a population rather than the variance of a sample. $\tilde{\sigma}^2/\sigma^2$

124. The symbol p without a tilde will be used to represent the proportion of favorable opinions in a **sample**. If you obtained a sample in which .3 of the opinions were favorable from a population in which .7 of the opinions were favorable, $\tilde{p}$ would equal ____ and p would equal ____. .7 .3

125. It will also be useful to have a symbol identifying the proportion of favorable opinions in a particular sample. For example, p_1 could represent the proportion of favorable opinions in the **first** sample. The symbol p_2 could represent the proportions of favorable opinions in the **second** sample. The symbol p_3 could represent the proportion of ______________ opinions in the ________ sample, and so on. favorable, thir

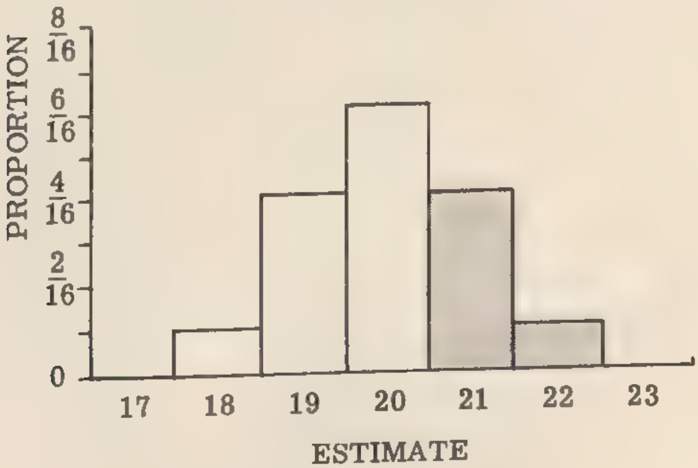
23. You could describe the variability of the estimates in B by saying that "the range of the estimates was ______." 6

In other words, the **range** equals ______ minus ______. 23, 17

24. The proportion of students who estimated that there were **more than** 21 marbles in the jar is simply the proportion who estimated that there were **22 or more**. If $\frac{2}{64}$ of the students estimated 22 marbles, $\frac{1}{64}$ of the students estimated 22 marbles, and none of the students estimated more than 23 marbles, you could say ______ of the students estimated there were more than 21 marbles in the jar. $\frac{3}{64}$

25. The proportion of students who estimated more than 20 marbles is simply the proportion of students who estimated there were 21 marbles, plus the proportion who estimated there were ______ marbles, plus the 22
proportion who estimated there were ______ marbles. 23

26. Imagine that the distribution of estimates was as follows: (Notice that we have shaded the columns representing subjects who estimated **more** than _____ marbles.) 20



117. Rather than obtaining a ____________ (biased/random) sample, in which some opinions in the population would have more opportunity to be included than would others, it was suggested that a ____________ (biased/random) sampling procedure be used, thereby providing every student's opinion the same opportunity to be included in the sample. — biased; random

118. One possible ____________ (random/biased) procedure for obtaining a sample would be to place each of the student opinions on a separate slip of paper, thoroughly mixing all of these slips of paper together in a basket, and drawing out as many opinions as you wished to include in the sample. — random

119. To determine how likely it was that a single sample obtained in this manner would be representative of the population, 100 samples were actually obtained by this procedure. Each of the 100 samples consisted of 5 opinions. These 5 opinions (the slips of paper on which the opinions were written) ____________ (were/were not) mixed again with the other slips of paper in the basket before the next sample was obtained (in order that all opinions would have the same opportunity of being included in each sample). — were

120. The slips of paper drawn for 1 sample had to be replaced before obtaining the next sample, since each opinion in the population had to have the ________ opportunity of being included in each sample if this was to be a random sampling procedure. — same

27. We pointed out earlier that the sum of all the proportions in a proportional (relative) frequency distribution has to equal _____. Since the height of each column represents the proportion of students who estimated a particular value, the total height of all the columns added together must equal _____.

1

1

28. According to the preceding graph, $\frac{1}{16}$ of the students estimated 18, $\frac{4}{16}$ estimated 19, _____ estimated 20, _____ estimated 21, whereas the remaining $\frac{1}{16}$ of the students estimated _____.

$\frac{6}{16}$ $\frac{4}{16}$

22

Thus, the total of the proportions represented by all the columns is _____/16, which equals _____.

16, 1

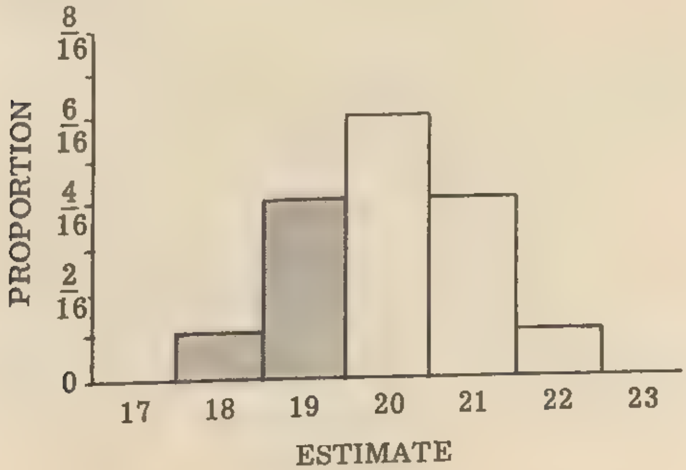
29. If we added together the heights of the two shaded columns in the previous distribution, they would form a column whose height was equal to the proportion of students who estimated **more than** _____ marbles. Notice that the proportion of students estimating **more than** 20 equals _____ plus _____, or $\frac{5}{16}$.

20

$\frac{4}{16}$ $\frac{1}{16}$

30. In the following distribution we have shaded the columns representing students who estimated _____ (fewer/more) than 20 marbles.

fewer



113. (Continued)

Only 5 samples of size 10 were obtained that would have led to an absolute error of estimate greater than .2, whereas _____ samples of size 10 were obtained that would have led to an absolute error greater than .2. 35

114. This illustrates a very important characteristic of random samples. The _____________ (larger/ smaller) the size of the samples, the less frequently would you expect to obtain samples leading to large absolute errors of estimation. larger

Another way of saying the same thing is to say that the larger the size of the sample, the more closely grouped will the sample statistics tend to be about the true parametric value. In other words, the larger the sample size, the _____________ (larger/ smaller) will tend to be the variability in your sample statistic. smaller

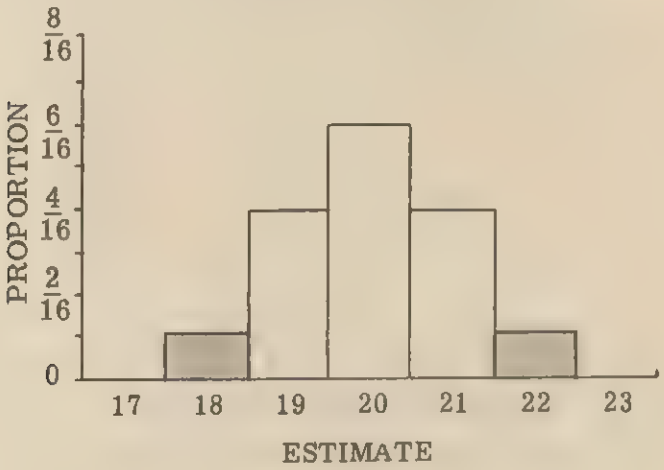
115. Let's review the important features of the previous illustrations. A complete collection of data was specified: the opinions (yes or no) of 10, 000 students at the university concerning the advisability of opening a foreign branch. This complete collection of data consisted, therefore, of ___________ observations of a variable we could call **opinion**. 10, 000

116. After this complete collection of data had been obtained, it was suggested that a satisfactory estimate of the student opinions might be made by simply obtaining **part** of the complete collection. In other words, instead of obtaining all ___________ (number) observations in the **population**, it might have been adequate to have merely obtained a s________ from the population. 10, 000 sample

30. (Continued)

A column as high as the two shaded columns combined would represent the proportion of students who estimated fewer than _____ marbles. This proportion would be ___. 20, $\frac{5}{16}$

31. In the following distribution we have shaded columns corresponding to people who estimated **fewer** than _____ 19
or **more** than _____ marbles. 21

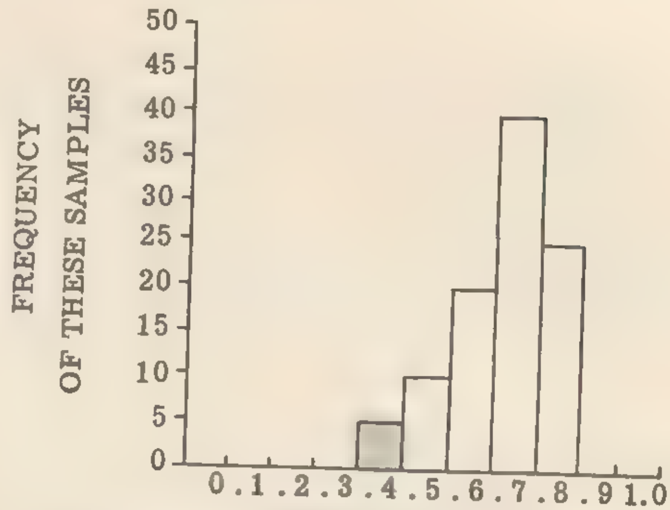


32. A subject who estimated 22 marbles would have made an **error** of 2 since there are actually 20 marbles in the jar. A subject who estimated 18 marbles would also have been in error by 2. The estimates indicated by the shaded columns in the previous distribution represent subjects who made an error of more than _____. 2

33. If the **mean** of the distribution of estimates was 20, an estimate of 22 would correspond to a positive deviation from the mean of _____. 2

Similarly an estimate of ____ would correspond to a negative deviation of -2. 18

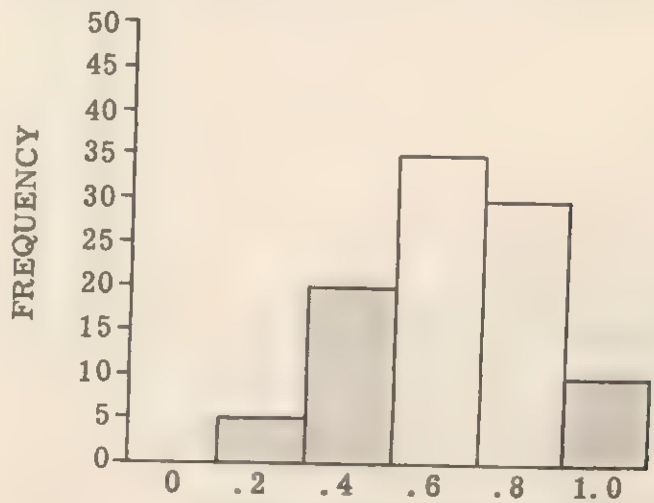
112. In the following graph, we have shaded those columns which represent samples that would lead to an absolute error of estimate **greater than** ____. (Remember, the population proportion was .7.) .2



SAMPLE STATISTIC

(PROPORTION OF FAVORABLE OPINIONS IN A SAMPLE OF TEN)

113. By increasing the size of the samples, we have apparently reduced the risk of obtaining a sample leading to an absolute error of estimate greater than .2. This fact is made most apparent by comparing the sampling distribution for samples of size 10 to the sampling distribution for samples of size 5 .



SAMPLE STATISTIC

(PROPORTION OF FAVORABLE OPINIONS IN SAMPLES OF FIVE)

34. The shaded columns in the previous distribution indicate the proportion of subjects whose estimates had either a positive deviation from the mean of ______ **or more**, or a negative deviation of _____ **or more**. (We use the words "or more" to indicate an estimate even farther away from the mean in either a positive or negative direction.)

2

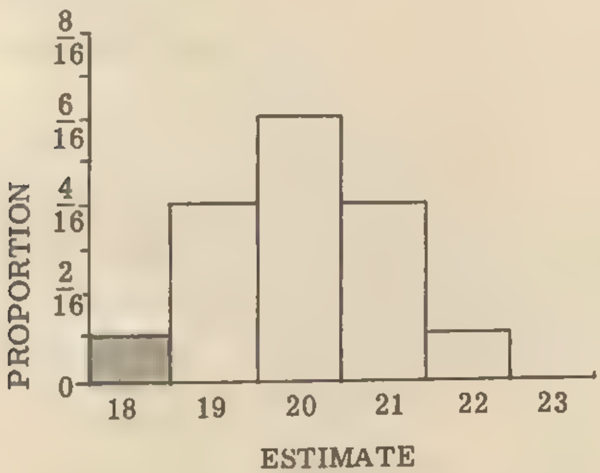
-2

35. The shaded column in the following distribution represents the proportion of estimates that had a ________________ deviation from the mean estimate of
positive/negative

______.

negative

-2

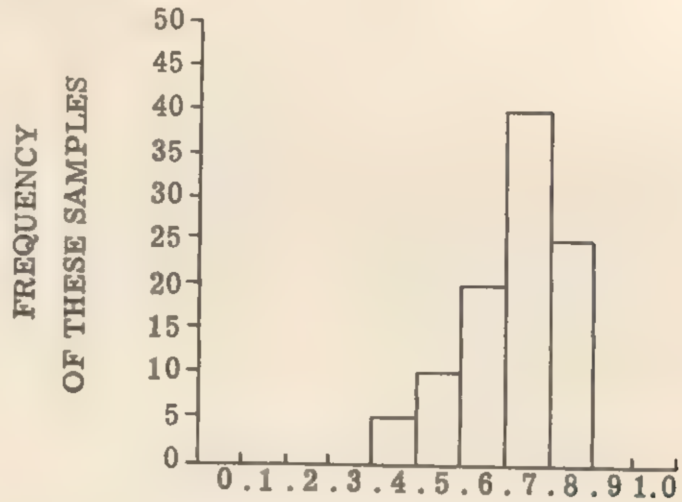


36. We will use the phrase **absolute deviation** or **absolute error** when we are interested only in the distance between a value and the mean and not in whether the deviation is positive or negative.

In other words, when we are interested only in the difference between a value and the mean and not in whether this difference is positive or negative, we will use the phrase __________ _____________ or _________ _________.

absolute deviation
absolute
error

110. This point is made clear by the following graph.



SAMPLE STATISTIC

(PROPORTION OF FAVORABLE OPINIONS IN A SAMPLE OF TEN)

This graph illustrates the same series of samples as is summarized in the previous table. Notice how most of the sample proportions are grouped around the true proportion of _____. .7

111. Sample proportions of 0, .1, .2, .3, .9, or 1 ______ occurred in the 100 samples. never

37. A value that deviates from the mean by +2 is the same **distance** away from the mean as a value having a deviation of -2. The value ten deviates from the value of eight by +2 (or simply 2). The value 6 deviates from 8 by -2. Both 6 and 10, however, are the same distance away from the value ______. This is what is meant when we say that the value 10 and the value 6 have the same **absolute deviation** from 8.

8

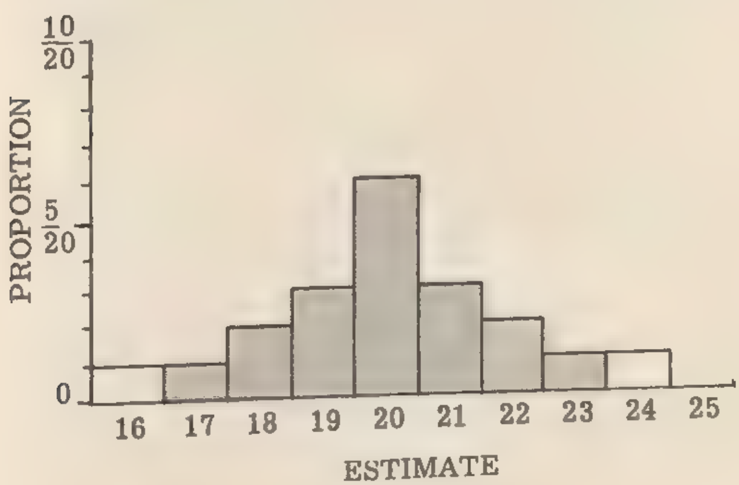
38. We just used an illustration in which 100 students attempted to estimate the number of marbles in a jar containing 20 marbles. An estimate of 22 would have been just as accurate as an estimate of 18, since both estimates would have been in error by ______.

2

39. If the mean of the estimates was 20, an estimate of 22 would represent a positive deviation of 2 and an estimate of 18 would represent a negative deviation of -2. Therefore, both an estimate of 22 and an estimate of 18 would represent the same __________ ____________ from the mean.

absolute deviation (error)

40. Suppose the distribution of student estimates was as follows:



104. Notice that samples in which exactly _______ of the 10 subjects had favorable opinions occurred more frequently than did any other type of sample. 7

105. Considering the **sample proportion** as an **estimate** of the population proportion, there would have been exactly ______ samples which would have led to a perfect estimate of the population proportion. 40

106. Sample proportions of ______, ______, or ________ would have led to an absolute estimation error of .1 **or less**. The frequencies of these three types of samples were ______, ______, and ______ respectively. .6, .7, .8 — 20, 40, 25

107. Adding these three frequencies together indicates that out of the 100 samples there were exactly ____ samples which would result in an absolute error of estimate of .1 or less. 85

108. Similarly, exactly ______ of the 100 samples would have led to an absolute error of estimate of .2 or less. 95

109. The only samples **actually obtained** that would have led to an absolute error of estimation **greater than** .2 were those samples in which _____ of the subjects had a favorable opinion. 4

Thus, _____ of the 100 samples would have resulted in a sample statistic of .4 and would thereby have led to an absolute error of estimation greater than ______. 5 — .2

40. (Continued)

The mean of this distribution is 20. Therefore, an estimate of 24 would represent a positive deviation of ______, and an estimate of 16 would represent a negative deviation of _____. 4 -4

41. On the previous distribution graph, the shaded columns represent estimates having an absolute deviation of **less** than _____. 4

42. The unshaded columns in the previous graph indicate the proportion of estimates having an absolute deviation from the mean of **more** than ______. 4

43. In a bell-shaped distribution, large absolute deviations are __________ (more/less) frequent than small absolute deviations. less

44. The **small** absolute deviations are **more** frequent because most of the observed values are clustered around the mean in a bell-shaped distribution. Estimates much larger or much smaller than the mean have a __________ (large/small) absolute deviation and are less frequent than those values with __________ (large/small) absolute deviations. large small

45. Students who estimated 25 marbles are **unusual** in the sense that few students made estimates that large. Similarly, students estimating only 15 marbles are also unusual, since there were very few estimates that small. In other words, the ______________ (larger/smaller) the absolute deviation of a student's estimate from the mean, the more **unusual** was his estimate. larger

102. If you actually carried out a sampling procedure of this sort, you would very likely obtain data similiar to that summarized in the following table.

Number of favorable opinions in a sample of 10	Proportion favorable opinions in the sample	Frequency of such samples
0	0	0
1	.1	0
2	.2	0
3	.3	0
4	.4	5
5	.5	10
6	.6	20
7	.7	40
8	.8	25
9	.9	0
10	1.0	0

The first column of the previous table indicates the 11 **possible** types of samples you could obtain, where each sample is identified by the number of favorable opinions it contains.

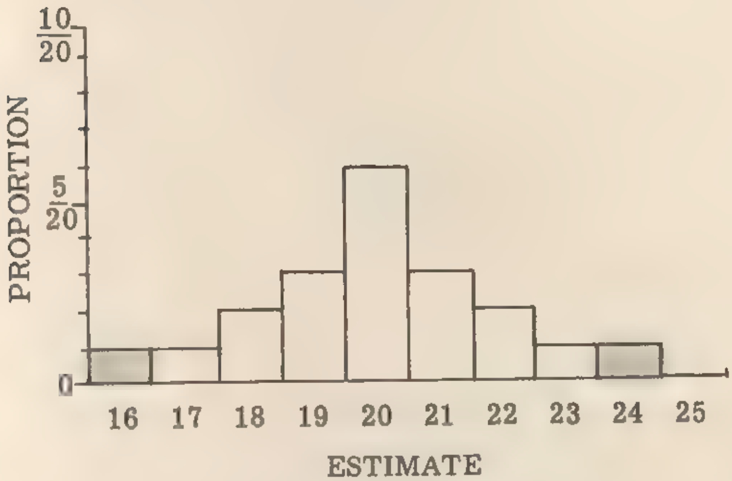
In the second column of the previous table, we have indicated the sample proportion corresponding to each of these 11 possible types of samples. In other words, if there were 7 favorable opinions in a sample of 10, the **proportion** of favorable opinions in that sample would equal ______. .7

7/.7

103. Finally, in the third column of the table, we have indicated how often each type of sample actually occurred in the 100 samples that were obtained. According to the table, there were exactly _____ samples in which all of the subjects had unfavorable opinions. 0

There were ____ samples out of the 100 samples in which exactly 4 of the 10 subjects had favorable opinions. 5

46. In the distribution shown below, we have shaded the columns representing the proportions of students who made either unusually ____________ or unusually ____________ estimates. large small



47. An estimate representing a large absolute deviation from the true number of marbles in the jar would be considered a __________ estimate. poor

good/poor

48. Suppose you conducted the marble estimation experiment with two groups of students, one group of 8-year-old students and another group of 18-year-old students. While you would expect the 18-year-old students to make some errors in their estimates, you would expect them to be more accurate than you would the younger students.

In other words, although the older students would make some mistakes, you would expect the absolute deviations of their estimates from the true number of marbles to be generally ______________ than those of the younger students. smaller

larger/smaller

99. The smallest number of favorable opinions you could have in any sample would be 0, whereas the largest number of favorable opinions would be ______. Thus, the 11 possible sample proportions would be: 10

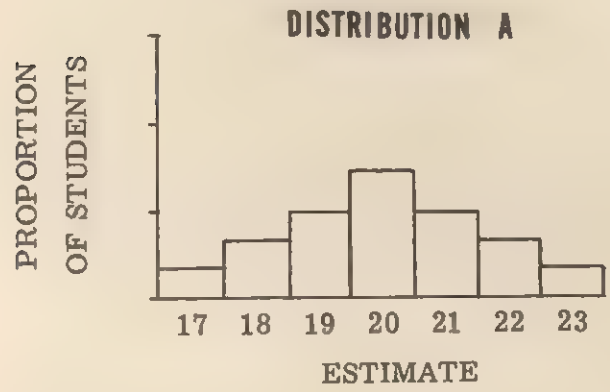
____	0
____	.1
____	.2
____	.3
____	.4
____	.5
____	.6
____	.7
____	.8
____	.9
____	1

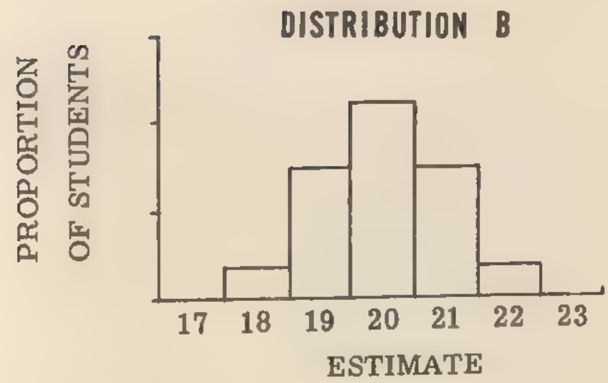
100. Thus, with a sample of 10, it ______________ (would/would not) be possible to obtain a sample statistic which was identical to the population statistic of .7. would

101. Exactly ____ out of the 10 opinions samples would have to be favorable in order for your estimate to be identical to the true population proportion. 7

48. (Continued)

Accordingly, if the following distributions represented the distributions of estimates from the two groups, Distribution ______ (A / B) is probably the distribution of estimates for the **younger** group. A





49. Distribution B represents the older students' estimates. It is clear that the older students tended to make more accurate estimates than did the younger students. The older students' estimates (Distribution B) tended to be closer to the true number of marbles than were the estimates of the younger students. You could say that the variability of the ______________ (younger/ older) students' estimates was greater than that of the other group. younger

95. (Continued)

population proportion. He might then suggest that instead of taking samples of ______ (number), as you did previously, you could take samples of 10. Even though it would be more time-consuming to obtain a sample of 10 than it would a sample of 5, it would still be easier than measuring the opinion of every person in the population. 5

96. Suppose you repeated the previous sampling procedure, this time obtaining samples of 10 rather than 5. In other words, you listed all of the students' names on slips of paper and placed these in a basket. After thoroughly mixing the slips in the basket, however, you would draw out ______ of these slips to form a sample of ______ rather than a sample of size 5. 10 10

97. Before obtaining the next sample, you would first replace the 10 slips you took out for the first sample.

Then, after once more mixing up the slips, you would take out another sample of ______. 10

98. In this manner, you could draw one sample of 10 after another, until you had exactly 100 samples, just as you did for the samples of 5. This time, however, your sample size is ______, whereas previously it was ____. 10, 5

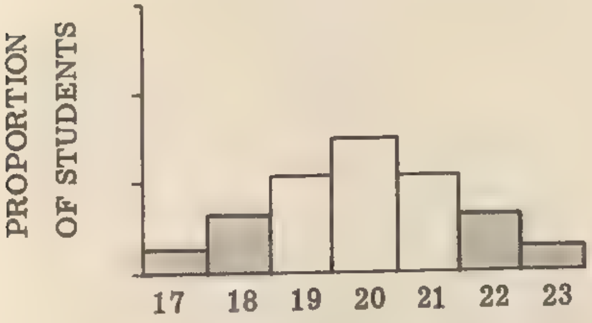
50. We could represent the difference in ____________ between the two distributions by the range. The range of Distribution A is ______ and the range of Distribution B is _____. variability

6

4

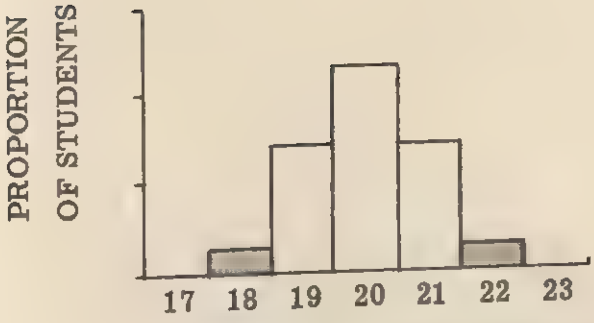
51. Estimates of either _____ or ______ would have an absolute deviation of 2 from the true number of marbles (20), since the first estimate would be 2 fewer than the actual value and the second estimate would be 2 greater than the actual value. 18, 22

52. The previous distributions are reproduced below. This time we have shaded the columns representing estimates whose absolute deviations from the true value of 20 were _____ or more. 2



YOUNGER STUDENTS' ESTIMATES

GRAPH A



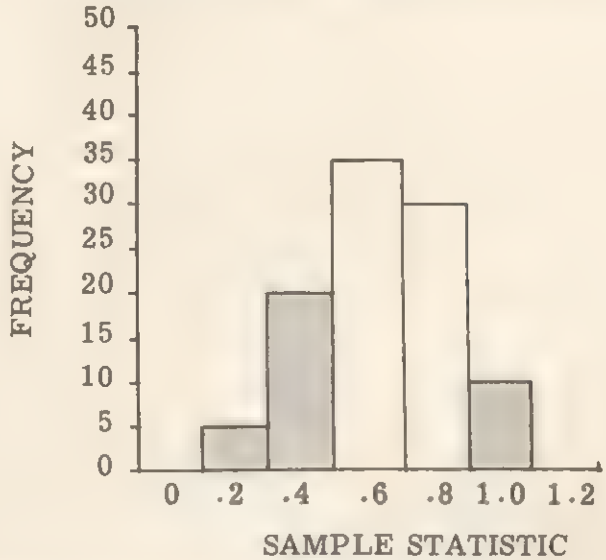
OLDER STUDENTS' ESTIMATES

GRAPH B

94. (Continued)

For example, look at the following graph. It is the same one we presented earlier; in this graph, however the ________________ columns represent the samples
shaded/unshaded
that would lead to an absolute estimation error **greater than** .1. (Remember, the population proportion was .7.)

shaded



(PROPORTION OF FAVORABLE OPINIONS IN SAMPLES OF FIVE)

95. While most of the samples of 5 you obtained would give you an estimate of the population proportion that differed from the true value by only .1, you still obtained a reasonable number of samples that would lead to larger error. You might be willing to accept the risk of making a poor estimate, since there was very little effort required in obtaining a sample of 5 compared to determining the attitude of every student in the population. You might tell your statistician friend, however, that even though it would be easier to ask some, rather than all, of the students their opinion, you are not willing to accept the risk of obtaining one of the relatively infrequent samples which would lead to a poor estimate of the

53. One way of comparing the **accuracy** of the younger and older students' estimates is to compare the proportion of students in each group whose estimates differed from the true value by 2 or more. The proportion of students who made errors **greater than one** is indicated by the _______________ columns in the previous graphs. shaded
shade/ unshaded

54. Estimates whose absolute deviation from the true value of 20 were greater than one occurred more often in the group of _____________ students than they did in the other younger
younger/ older
group.

55. Among the younger students, an estimate of 22 was __________ **unusual** than an estimate of 21, since the more
more/ less
proportion of younger students who estimated 22 was _______________ than the proportion of younger students smaller
greater/ smaller
who estimated 21.

56. An estimate of 22, however, was __________ **unusual** for less
more/ less
younger students than it was for older students, since the proportion of younger students who made estimates of 22 is _______________ than the proportion of older larger
larger/ smaller
students who made estimates of 22.

57. Compared only with the rest of the people **in his own age group**, a young student who estimated 22 would not have performed quite so poorly as an older student who had estimated 22, since an estimate of 22 was more common (occurred more frequently) in the _____________ group younger
younger/ older
than it was in the other group.

91. We shall use the phrase **absolute estimation error** to represent the difference between your estimate and the true parametric values, **disregarding which of the two was larger**. In other words, both a sample proportion of .6 and a sample proportion of .8 would lead to an **absolute** estimation error of .1. Since $\tilde{p}$ = .7 both sample proportions differ from the true value by .1, even though the sample proportion of .6 is ________ than the true value by .1 and the sample proportion of .8 is ________ than the true value by .1. — smaller / larger

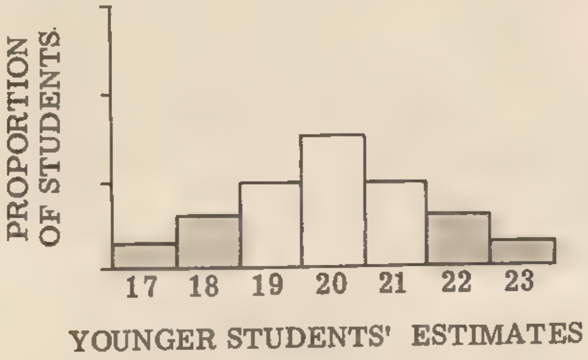
92. It is perfectly clear from the previous table that large absolute estimation errors are ________ (less/ more) frequent than small estimation errors. — less

93. In other words, large absolute estimation errors (whether they be the result of overestimating the population parameter or of underestimating the population parameter) are **unusual** compared to estimates that are closer to the population parameter. In fact, an absolute estimation error of .5 occurred _____ times out of the 100 samples. — 0

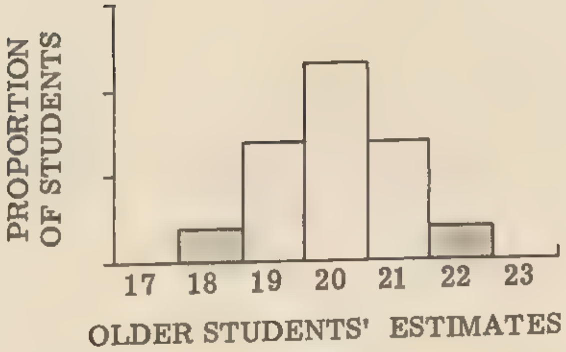
94. To review, at the statistician's suggestion you considered the sampling distribution of samples of 5 opinions, each collected by a random procedure from a population of 10, 000. The **experimental** sampling distribution you obtained suggested that most samples obtained in this way would probably give you a fairly good estimate of the population proportion. However, there were an appreciable number of samples which would have led to an absolute error greater than .1.

58. Suppose you decided to give a prize to all the students whose estimates were within one marble of the true value. Or, to put it differently, suppose you decided to give a prize to all the students whose estimates had an absolute deviation from the true value of _____ or less. 1

59. The _______________ columns in the following graphs
shaded/ unshaded
indicate the proportions of students in each age group who would receive a prize. unshaded



GRAPH A



GRAPH B

87. The following table contains the same information as did the previous graph. In additioh, we have indicated the **error** of **estimate** which would occur for each of the six possible kinds of samples. A sample in which none of the opinions were in favor would lead to an estimate of the population value that would be in error by -.7.

Frequency of Opinions in Favor	Frequency of Such Samples	Error of Estimate
0	0	-.7
.2	5	-.5
.4	20	-.3
.6	35	-.1
.8	30	.1
1.0	10	.3

Samples in which only .2 of the 5 opinions sampled (1 out of 5) were in favor would lead to an error of estimate equal to ______, since .2 minus .7 equals ___. -.5, -.5

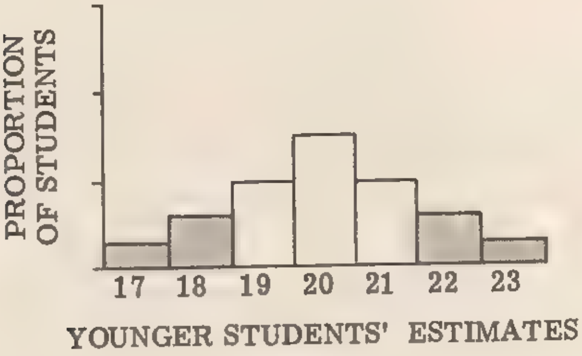
88. Proportions in which 3 out of the 5 opinions were favorable give you a sample proportion equal to _____. .6
This sample proportion would lead you to make an estimate which was in error by ______. -.1

89. Samples with 4 of the 5 opinions in favor give a sample proportion equal to .8. Since .8 is greater than the parametric proportion, those samples would give you an estimate which would be in error by _____. .1

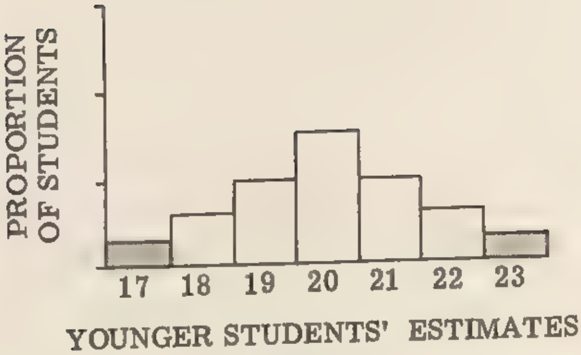
90. If all 5 of your sample opinions were favorable, your estimate of the population proportion would be _____. 1
This would represent an error of estimate of _____. .3

60. If you only gave a prize to those younger students whose estimates were within 1 of the actual value, the unshaded columns in Graph ______ (below) would indicate A
A / B

the proportion of younger students who received a prize.



GRAPH A



GRAPH B

61. Since the older students were more accurate in estimating the number of marbles, ______ of their more
fewer/ more

estimates would meet the requirements for a prize than would those of the younger students because ______. more
more/ less

of the older students had estimates close enough to the true value.

82. According to the previous sampling distribution, it ______________ (would/would not) be possible to obtain a sample with .2 of the opinions in favor, but it ______________ (would/would not) be unusual to obtain a sample of this sort. — would; would

83. Of the 100 samples (of 5 opinions each) that were obtained, there were only _____ samples which would have led you to make an estimate that differed by **more than** .3 from the true population proportion. — 5

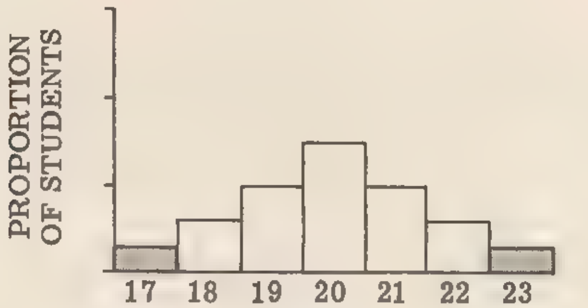
84. Since the population proportion was .7, only proportions less than _______ would differ from this true proportion by more than .3 — .4

85. If you used a sample statistic of .4 as your estimate of the population statistic, the difference between your estimate and the population proportion would be .7 minus .4, which would equal .3. The question was: What sample statistic would differ from the true population value by **more than** .3? The only possible values of the sample statistic (in samples of 5) that would lead to estimates differing by **more than** .3 from the true value are samples having **no** favorable opinions or **one** favorable opinion. In other words, sample proportions of ______ or ______ would differ from the true population proportion of .7 by more than .3. — 0, .2

86. According to the previous graph, there were no samples in which all the subjects were opposed to the proposal, and there were only ______ samples in which one of the subjects was in favor of the proposal. — 5

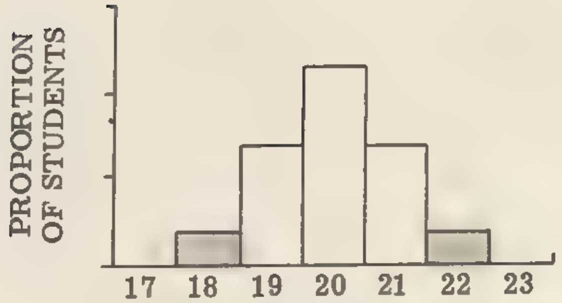
62. The unshaded columns in the following graphs indicate the proportion of older students who would receive a prize if we only gave a prize to students who made errors of ______ or less. 1

The other graph indicates the proportion of younger students would receive a prize if we gave a prize to all younger students who made an error of ______ or less in their estimates. 2



YOUNGER STUDENTS' ESTIMATES

GRAPH A



OLDER STUDENTS' ESTIMATES

GRAPH B

78. The preceding sampling distribution suggests that samples obtained in this manner tend to have sample statistics quite similar to the population statistic. This is obvious on the previous graph, since most of the samples had proportions equal to .6 or .8, whereas samples with proportions equal to .2 or 1 were much ________ (more/less) frequent. — less

79. Note that the sampling distribution is roughly bell-shaped, with most of the sample proportions clustered around the true population value of ________ (.4/.7). — .7

80. This sampling distribution suggests that if you used the same sampling procedure to obtain a single sample and if you used the sample statistic to estimate the population statistic, you would run little risk of making an estimate of the population proportion equal to .2, since a sample proportion equal to .2 was quite **unusual** (very ________ (few/many) of the samples you obtained had a proportion equal to .2). — few

81. Samples with proportions equal to .6 or .8 were not unusual. In fact, they were quite typical, since the frequencies of samples in which .6 or .8 of the students were in favor of the overseas campus are relatively ________ (high/low) compared to the other possible values of the sample statistic. — high

62. (Continued)

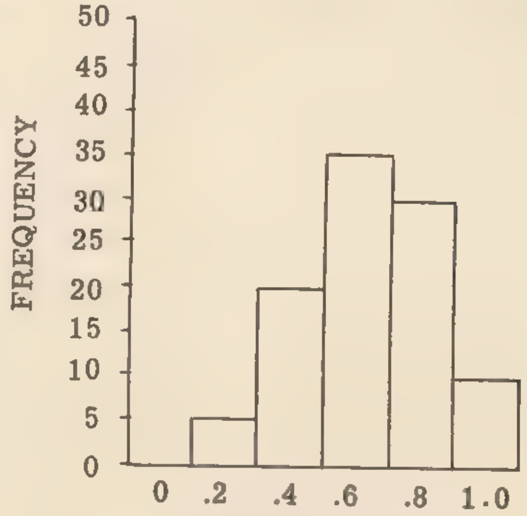
The proportion of younger students obtaining prizes and the proportion of older students obtaining prizes _______________ be approximately the same if we used
would/would not
these two rules for awarding prizes. would

63. We could conclude from these considerations that a _______________ student whose estimate was within 1 younger
younger/older
of the actual value was doing just about as well **in relation to the rest of the people in his age group** as was a(n) _______________ student whose estimate was within 2 older
younger/older
of the actual value.

64. By considering the difference in variability between the two groups, we established rules for giving prizes whereby approximately the same proportion of students received prizes in each group. Since large errors (absolute deviations) were more frequent in the younger group, a _______________ student would not be required to be younger
younger/older
quite as accurate in order to win a prize as would a student in the other age group.

65. The preceding example illustrates why it is often useful to consider the **variability** of a distribution when you are evaluating a particular observed value. For example, imagine you were teaching a course in psychology. You gave your students two examinations during the semester. Suppose there were 10 questions on each test and the student received either a score of 1 or a score of 0 on each question. The possible total score on each test would be somewhere between ___ and ___. 0 10

75. The previous table is similar to a frequency distribution table. It is also possible to draw a frequency distribution graph (similar to the ones we have seen earlier) indicating the distribution of these sample statistics.



SAMPLE STATISTIC
(PROPORTION OF FAVORABLE OPINIONS IN SAMPLES OF FIVE)

Notice that there were more samples in which the proportion of students in favor of the proposal equaled ______ than any other possible value of the sample statistic. .6

76. The **highest** column on the graph represents the frequency of samples in which the sample statistic (proportion of opinions in favor) equaled ________. .6

77. The distribution shown in both the previous table and the previous graph is that of a sample statistic. Thus, the distribution could be described as a s______________ distribution. sampling

65. (Continued)

Suppose the results of these tests were those shown in the following frequency tables.

Examination A

Score	Frequency
0	0
1	0
2	0
3	1
4	4
5	5
6	5
7	4
8	1
9	0
10	0

Examination B

Score	Frequency
0	0
1	0
2	1
3	2
4	3
5	4
6	4
7	3
8	2
9	1
10	0

According to these data, there were _____ students in the class. 20

66. There must have been 20 students in the class because the sum of the ________________ in each frequency table is 20. frequencies

67. The **variability** of the scores on Test ______ (A / B) appears to be slightly larger than the variability on the other test. A

The **range** is one statistic which would represent this difference in variability, since the range was ______ on Test A and _____ on Test B. 5 7

68. It also appears that σ^2 was larger on Test _____ (A / B) than on the other test. A

72. (Continued)

Notice that this table is very similar to a frequency table. There are ____ possible values of a variable called "proportion of opinions in favor." 6

The frequencies in the second column of the table indicate how many samples with each possible value of the sample statistic were obtained. The sum of these frequencies indicates that exactly ______ samples were obtained. 100

73. Notice that there were **no** samples in which all of the students were opposed to the overseas campus. However, there were _____ samples in which all of the students were in favor of the overseas campus. 10

74. Since there were 5 opinions in each sample, it is impossible to obtain a sample proportion equal to the true population proportion of .7. However, in 35 of the samples there were exactly _____ students in favor of the proposal, and in 30 of the samples exactly _____ of the students were in favor of the proposal. In other words, 35 samples had a sample proportion of _____ and 30 samples had a sample proportion of ______.

Answers: 3, 4, .6, .8

69. The preceding absolute frequency distributions could be converted to **relative** frequency distributions by dividing each frequency by _______ in order to convert it to a _______________. 20 proportion

The two proportional distributions which you would obtain in this manner are shown in the following tables.

Examination A

Score	Frequency
0	0
1	0
2	0
3	1/20
4	4/20
5	5/20
6	5/20
7	4/20
8	1/20
9	0
10	0

Examination B

Score	Frequency
0	0
1	0
2	1/20
3	2/20
4	3/20
5	4/20
6	4/20
7	3/20
8	2/20
9	1/20
10	0

Notice that $\frac{1}{20}$ of the students received scores of **3 or lower** on Examination A, whereas ______/20 of the students received scores of **3 or lower** on Examination B. 3

70. The **mean** of both previous distributions (Examination A and B) is $5\frac{1}{2}$. A score of 5 would have a negative deviation of $-\frac{1}{2}$ from the mean and a score of 6 would have a positive deviation of _____ from the mean. $\frac{1}{2}$

71. The only two scores which have an absolute deviation of $\frac{1}{2}$ or less (from the mean of $5\frac{1}{2}$) are _____ and _____. 5, 6

72. One-half $(\frac{10}{20})$ of all 20 students received scores whose absolute deviation from the mean was less than $\frac{1}{2}$ on Examination _______. A

A / B

69. Suppose you agreed to this demonstration and, with the aid of your secretary, made the necessary arrangements to obtain the samples in this manner. If you decided to obtain samples of 5 opinions, you would place the slips of paper indicating student opinions in a large basket, mix the basket thoroughly, and draw out _____ slips of paper.

5

70. Before proceeding to obtain the next sample, you would have to replace the first 5 slips of paper you had taken out. Otherwise, the students whose opinions were represented in the first sample _______________ (would/ would not) have an equal opportunity of being included in the second sample.

would not

In other words, in order to give each student's opinion the same opportunity of occurring in each sample, there would have to be ___________ slips of paper in the basket each time a sample was obtained.

10, 000

71. The number of students in favor of the overseas campus in each sample could be anywhere from 0 to 5. Thus, the **proportion** of students in each sample who were in favor of the overseas campus would have to equal **0**, **.2**, ____, ____, ____, or ____.

.4, .6, .8, 1

72. Imagine that 100 samples consisting of 5 opinions each were obtained in this manner. The results could be summarized in the following table.

Proportion of Opinions in Favor	Frequency of Such Samples
0	0
.2	5
.4	20
.6	35
.8	30
1.0	10

73. The scores 5 and 6 are the only scores deviating from the mean by $\frac{1}{2}$ or less on both examinations. The proportion of students who received scores of 5 or 6 is equal to

$$\frac{5}{20} + \frac{5}{20} = \frac{10}{20}$$

on Examination ______ (A / B). A

74. The proportion of students who received **either** a score of 5 **or** 6 on Examination B is equal to ______ + ______, which equals ______. $\frac{4}{20}$, $\frac{4}{20}$ $\frac{8}{20}$

75. It was slightly more **unusual** for a subject's score to have an absolute deviation from the mean of more than one-half on Examination ______ (A / B) than on the other A

examination, since a higher proportion of students received grades of 5 or 6 on Examination ______ (A / B). A

76. The absolute deviation of an observed value from the mean does not necessarily indicate how **unusual** that value was in relation to the rest of the distribution. If most of the values in the distribution were grouped very closely around the mean, the proportion of values whose absolute deviation from the mean was greater than 1 might be very small. On the other hand, if the distribution were quite variable, a much higher proportion of the values might have absolute deviations from the mean greater than 1.

66. While you might be unfortunate enough to obtain a sample in which only 1 out of the 10 (.1) of the subjects were in favor of the overseas campus, you would probably find that such highly unrepresentative samples occurred infrequently. In other words, an unrepresentative sample would be __________ unusual than a representative sample. more

more/less

67. The time and effort you could save by taking a sample rather than a complete collection of data could be balanced against the **risk** of obtaining a highly unrepresentative sample. If the risk of making a large error of estimate were **small** compared to the amount of effort that you would save by only having to collect a sample, it ______________ be more desirable to use the sampling procedure. would

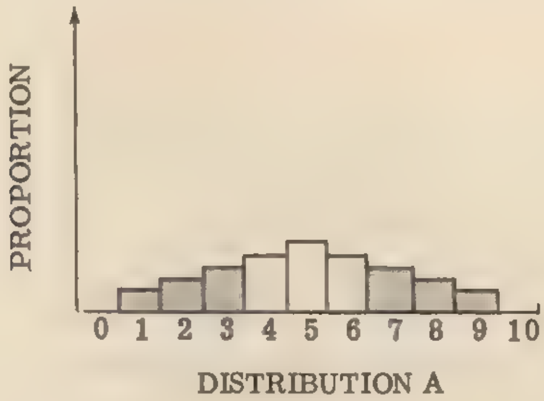
would/would not

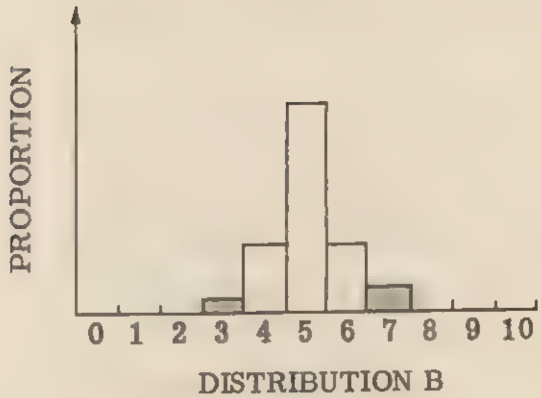
68. Let's suppose the statistician suggested the following procedure as a means of clarifying his proposal. Since the data which defined the population of student opinion had already been collected, it would be possible to obtain as many samples of this population as you desired. Instead of assigning each student a number from 1 to 10,000, you could actually write each student's opinion on a piece of paper, place all the papers in a basket, shake the basket well, and draw out a sample of however many opinions you wished. This would be a way of obtaining a ________________ sample from the population of student opinions. random

random/biased

76. (Continued)

For example, consider the two distributions shown below.





The mean of both distributions is 5. However, the variability of Distribution ______ appears to be greater than the variability of the other distribution. A

A / B

77. While the mean of both distributions is 5, absolute deviations greater than 1 occur much more frequently in Distribution ______ than they do in the other distribution. A

A / B

64. The statistician might propose a **random sampling** procedure of the following sort. Every student enrolled at the university would be assigned a different number from 1 to 10, 000. Each of these numbers would be written on a separate slip of paper and all of these slips could be placed in a large basket. Then, the basket would be thoroughly shaken so that all the slips of paper inside were scrambled. If you wanted to obtain a sample consisting of 10 students' opinions, for example, you would simply draw out 10 slips of paper identifying 10 students. You could obtain these 10 students' opinions by contacting them directly. Their answers would represent a sample of 10 student opinions. This would be a ____________ (random/biased) procedure for obtaining a sample of 10 opinions, since every student's opinion had the ____________ opportunity to be included in the sample.

random

same

65. You might then point out to the statistician that even a procedure of this sort would not preclude the possibility of obtaining an **unrepresentative** sample. For instance, you might be unlucky and obtain a sample of 10 opinions of which 9 were against the overseas campus and 1 was in favor of the overseas campus.

If .7 of the whole student body were actually in favor of the overseas campus, the population proportion would be .7, whereas the corresponding sample statistic would be _______. This would mean your **error of estimate** would be .1 minus .7, which equals _____.

.1

-.6

78. While the value 7 represents a deviation from the mean of ______ in either distribution, values with deviations that large or larger were more unusual in Distribution ______ than they were in the other distribution. 2

A / B B

79. Statisticians describe an observed value in a way that takes into account the variability of the distribution. You saw earlier how an observed value could be represented by its deviation from the mean rather than by its actual value. In addition, it is sometimes useful to indicate the relationship between that deviation and the standard deviation of the distribution. For example, if the mean of a particular distribution were 10, the value 15 would have a deviation from the mean of _____. If the standard deviation of the distribution were 5, you could say the value 15 deviates from the mean by exactly **one standard deviation**. Similiarly, since the value 20 has a deviation from the mean of 10, 20 would be exactly __________ standard deviations away from the mean if $\sigma = 5$. 5 two

two/three

80. The variance of a distribution is simply the typical (**mean**) squared deviation from the mean of that distribution. The **standard deviation** squared would equal the ________________. variance

61. (Continued)

sample of the students' opinions. In other words, he was suggesting that you might have viewed the 10,000 opinions as a __________ and obtained a __________ consisting of part of this complete collection of 10,000 opinions. Suppose he suggested that you could have used the proportion of favorable opinions in your **sample** as an estimate of the actual proportion of favorable opinions in the population. Or, to put it more exactly, you could have used a __________ (population/sample) statistic to estimate a __________ (population/sample) statistic.

population, sample

sample

population

62. You might be concerned with exactly how the sample should be obtained. If it were obtained improperly, it could be very misleading. For example, you might point out to him that the sample was obtained by questioning all the people in the French department, where many people could be expected to favor spending time in a foreign country, in which case you would more than likely __________ (overestimate/underestimate) the population proportion of students in favor of an overseas branch.

overestimate

63. The statistician would probably answer that you were describing a procedure for collecting samples which would give some students more opportunity of being included in the sample than it would other students. The statistician would suggest using a sampling procedure that gave every student's opinion an equal opportunity of being included in the sample. In other words, he would say that rather than using a __________ (biased/random) sampling procedure, you should use a __________ (biased/random) sampling procedure.

biased

random

81. Consider the following relative frequency distribution.

VALUE	PROPORTION
2	1/20
3	4/20
4	10/20
5	4/20
6	1/20

It can be shown that the variance of this distribution is 16. Therefore, the ________ __________ is equal to $\sqrt{16}$ or 4. standard deviation

82. The mean is also 4. Thus, the values 3 and 5 ________ (are/are not) more than one standard deviation away from the mean because the absolute deviation from the mean of 4 for both the value 3 and the value 5 is ______. are not 1

83. You could say the value 5 is $\frac{1}{4}$ of a standard deviation away from the mean, since the value 5 represents a deviation from the mean of 1 and 1 is $\frac{1}{4}$ the size of the standard deviation 4. Similiarly, the value 6 would be ________ ($\frac{1}{4}$ / $\frac{1}{2}$) of a standard deviation from the mean, since the value 6 deviates from the mean by 1 and 1 is ______ the size of the standard deviation 4. $\frac{1}{4}$ $\frac{1}{4}$

84. We have described the distance or difference between a particular value and its mean as the ______________ of that value from that mean. deviation

59. (Continued)

One of the things you might wish to take into account before making your decision would be the opinions of the students themselves. Therefore, let's suppose that at the beginning of the school year the pertinent question was printed on all registration forms, a question which the student was required to answer **yes** or **no**. This question simply asked whether or not the student thought an overseas campus would be a valuable addition to the facilities for education offered by the university. Suppose it was determined in this manner that exactly .7 of the whole student body responded favorably to the idea of an overseas campus. If there were 10, 000 students at the university, your data would consist of
__________ observations of a variable you could call "**opinion**," whose two possible values were __________ and __________.

10, 000
yes
no

60. If .7 of the students were in favor of the overseas branch, the frequency of "yes" answers divided by __________ equals .7 , and the frequency of "no" answers divided by ________ equals _______.

10, 000
10, 000, .3

61. Suppose that after you had taken this survey, you had lunch at the faculty club with a professor from the Statistics department. When you explained the procedure you used for determining student opinion, he proposed what he felt would be a simpler and faster means of assessing student opinion. Instead of taking the time to determine every student's opinion, he suggested it might have been adequate to have simply obtained a

84. (Continued)

If a value is smaller than the mean, you say that it has a ________________ (positive/negative) deviation. If a value is larger than the mean, that value has a ________________ (positive/negative) deviation.

negative

positive

If the distribution had a mean of 6, the value 7 would represent a deviation of _______ and the value 5 would represent a deviation of ______.

1

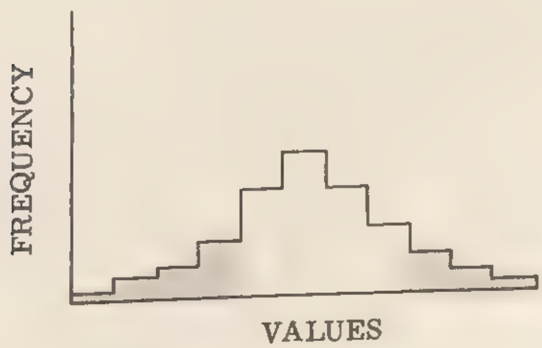
-1

85. Suppose the distribution of your data were approximately "bell shaped." You would know that most of the observed values were clustered around the mean with progressively fewer and fewer observed values farther away from the mean. In other words, the frequency of values having a deviation of 2 is probably greater than the frequency of values with a deviation of ______ (3 / 1).

3

86. The values farthest away from the mean in a "bell-shaped" distribution are often referred to as the **tails** of the distribution (since the distribution appears to taper into a tail at the extreme values). For example, the ________________ (shaded/unshaded) areas in the following distribution would be referred to as the tails of the distribution.

shaded



56. Suppose you were the principal of a school and you had selected ten of your best mathematics students to compete in a contest with another school. The grades in mathematics of the ten students ________ (would/would not) be a **random** sample from a population consisting of the mathematics grades of all the students in the high school. — would not

57. The grades scored by these ten students would **not** be a random sample from the population of mathematics grades, since each of the grades in the population ________ (did/did not) have the same opportunity of being included in the sample. Because of the way the sample was formed, students with good grades had ________ (the same/a better) chance of being included than did those with poor grades. — did not; a better

58. Later you will see how it is often possible to **predict** the distribution of sample statistics (i.e., the sampling distribution) if samples are drawn in a **random** or unbiased manner. First, however, let's consider how knowledge of a sampling distribution would help you to make estimates of population statistics from sample statistics. In other words, let's consider how knowledge of the sampling distribution would help you to **estimate** statistics that describe the ________ on the basis of statistics describing the ________. — population; sample

59. Suppose you were the president of a large Midwestern university who was trying to decide whether or not to open an overseas branch of the university. The object of having an overseas branch would be to provide students with an opportunity for living and studying in a foreign country sometime during their period of attendance at the university.

87. The values from the ______________ absolute largest
largest/smallest

deviations from the mean appear in the **tails** of a bell-shaped distribution.

88. In a so-called "bell-shaped" distribution, the farther a value is from the mean (the more it deviates from the mean), the ______________ will be its frequency. smaller
larger/smaller

89. You should recall that the **standard deviation** (σ) is a deviation which, squared, would equal the ______________ . variance

If the **variance** of a distribution were 9, a deviation of 4 would be ___________ than one **standard deviation**. greater
greater/less

90. If the mean of your distribution were 10 and the standard deviation of the distribution were 4, a score of 14 would have a deviation from the mean equal to _____ standard 1
1 / 2

deviation(s).

91. Another way of indicating that a particular value deviates from the mean by one standard deviation is to say that that value equals a **standard score** of one. Thus, if the value 14 represents a **standard score** of one in a distribution with a mean of 10, the standard deviation of the distribution must be ______. 4

92. If the distribution had a standard deviation of 2 and a mean of 10, then the value ______ would deviate from 12
12/14

the mean by one standard deviation. Therefore, 12 would represent a s___________ s___________ of one. standard score

53. Of the two statements shown below, Statement ______ (A/B) would more properly describe the procedure by which samples were obtained in the preceding illustration. B

Statement A:

Every yearly income in the population had the same opportunity of occurring in each sample as did any other.

Statement B:

Certain yearly incomes in the population were more likely to occur in samples taken in one part of town than in the samples taken in the other part of town. ______ B

54. A procedure for selecting a sample from a population that gives some members of the population a greater opportunity of being selected than other members is called a **biased** sampling procedure. A sampling procedure in which all members of the population have the **same** opportunity of being selected is called an **unbiased** or **random** sampling procedure. The sampling procedure used in the previous illustration is an example of a(n) ______________ (biased/random) sampling procedure. biased

55. Consider a population consisting of 100 observations. Suppose you listed each observation on a card and then shuffled the cards very thoroughly. If you then picked ten of these cards, each of the 100 observations would have had the same opportunity of being included in the sample. Therefore, this method of selecting a sample would be an example of a ____________ sampling procedure. random

93. If a particular value were said to equal a **standard score** of 2, that value would have a deviation from the mean equal to twice the standard deviation of the distribution. If a particular value equaled a standard score of -2, that value would represent a **negative deviation** from the mean equal in size to twice the ____________ ____________. standard deviation

94. Below are a list of values forming a distribution whose mean is 10 and whose variance is 25. Since the **variance** is 25, the size of a **standard deviation** is _____. $5 = \sqrt{25}$

VALUE	DEVIATION FROM 10	STANDARD SCORE
15	5	·1
5	-5	-1
10	0	[illegible]

95. Notice that the first value, 15, deviates from the mean of 10 by ______. Since the standard deviation of the distribution is 5, a deviation of 5 would be the same size as one standard deviation. This is all we mean when we indicate (as we did in the previous table) that the value 15 is equal to a **standard score** of _____. 5 1

96. The value 5 in the previous distribution represents a negative deviation from the mean of -5 and, therefore, is equivalent to a standard score of _____. -1

1 / -1

49. The most common or frequently occurring sample mean in the distribution shown on Graph A (above) is a mean of $________ per year, while the modal value of the sample means on Graph B is $________ per year.

$12, 000
$5, 000

50. It would be reasonable to suppose that the distribution of twenty samples shown in Graph _____ (A/B) is based on the samples collected in the lobby of the expensive hotel, while the distribution of sample means shown in Graph _____ (A/B) is more likely that of the twenty samples collected in the poor section of town.

A
B

51. The distribution of sample means shown on Graph A were probably based on the samples taken in the expensive hotel (rather than in the poorest section in town), because the sample means tend to be ________ (larger/smaller) on Graph A than they are on Graph B. In other words, you would expect to meet more people with large incomes in the expensive hotel than you would in the poor section of town.

larger

52. Both the twenty samples taken in the expensive hotel and the twenty samples taken in the poor section of town ________ (can/cannot) be considered as samples from a population consisting of the yearly incomes of all the people in the city. Although all the samples were from the same population, you were more likely to sample a person with a ________ (high/low) yearly income in the expensive hotel than you would in the poor section of town.

can
high

146. Most of the values in a normal distribution are clustered around the mean, with fewer and fewer values farther away from the mean (i.e., the distribution is "bell-shaped"). Thus, values representing a Z-score larger than 2 would be ________ (more/less) frequent than values representing Z-scores less than 2. less

147. You have already seen that .95 of all the values in a normal distribution are within two standard deviations of the mean. Therefore, .95 of all the values in a normal distribution would represent Z-scores between -2 and ______. 2

148. Thus, a Z-score is a useful way of representing values in a normal distribution since it indicates how unusual such values are, regardless of the mean or variance of the distribution. No matter what the mean or variance of the normal distribution, you would know that a Z-score as large as 1 was ________ (more/less) frequent than a Z-score as large as 3. more

149. Suppose you were told that your score on the last Psychology examination was approximately equal to a Z-score of two. This would imply that the distribution of test scores was approximately normal and that your score was about ________ standard deviations above the mean of the distribution. two

150. Furthermore, since only .05 of the values in a normal distribution are farther than two standard deviations from the mean, and since these extreme values are divided equally between negative and positive deviations, you would know that only about ________ (.05/.025) of the students had made a ____________ (better/poorer) grade on the test. .025 better

142. (Continued)

a collection of data **as if** it had a **normal distribution,** even though any actual collection of data ________ (can/cannot) have a perfect normal distribution. — cannot

Suppose you told someone that your data had an approximately normal distribution. He would know that approximately .95 of all the values in the distribution were within two standard deviations of the mean. If a particular observed value represented a **standard score** of 2, he would know that only about ________ (.05 / .01) of the observed values were farther away from the mean of your data than was that particular value. — .05

143. A **standard score** from a normal distribution is often called a Z -score. If a value from a normal distribution represented a Z -score of -2, you would know the value was two standard deviations ________ (above/below) the mean. — below

144. A value would represent a Z-score of 2 if it were from a **normal distribution** and had a ________ (positive/negative) deviation from the mean equal to ____ standard deviations. — positive; 2

145. Since a standard score is simply the number of standard deviations between a value and the mean, a Z-score is simply the ________ of standard deviations between a value and the mean in a ________ ________. — number; normal distribution

Turn the book around and continue with frame 146 on page 240.

151. To summarize then, a standard score (or Z-scores if the distribution is normal) are a convenient way of representing a value, since it indicates how many _____________ _______________ that value is away from the mean.

standard deviations

152. This is particularly useful to you in the case of a **normal distribution**, since you know exactly what _____________ of the values in the distribution are within any particular number of standard deviations from the mean.

proportion

140. By imagining the proper dotted lines, you could determine that approximately ______ (.36/.46) of all the values in a normal distribution were within 1/2 a standard deviation of the mean. .36

141. If .36 of the values were within 1/2 a standard deviation of the mean, the remaining values (1 - .36 = .64) would be **1/2 or more** standard deviations from the mean. Since the normal distribution is **symmetrical**, these .64 of the values would be equally divided between each tail of the distribution. Thus, there would be ______ of the values in one tail and ______ of the values in the other tail. .32 .32

142. The normal distribution is referred to as a **theoretical distribution** because it is defined in terms of a mathematical equation or formula rather than in terms of an actual collection of data. An actual collection of data may have a distribution similar to a normal distribution. However it can never have a distribution that is exactly normal, in the same sense that an actual circle can never be **perfectly round** since it would always be possible to demonstrate some slight difference between the actual object and the mathematically defined **perfect** circle. It is often useful, however, to treat circles as if they were perfectly round, even when, in fact, they are not.

A doughnut ______ (is/is not) **exactly** round, but a person who packaged doughnuts would probably find it convenient to treat them as if they were round, since the difference between their shape and a perfect circle is not important to him. In the same sense, it is often useful to treat is not

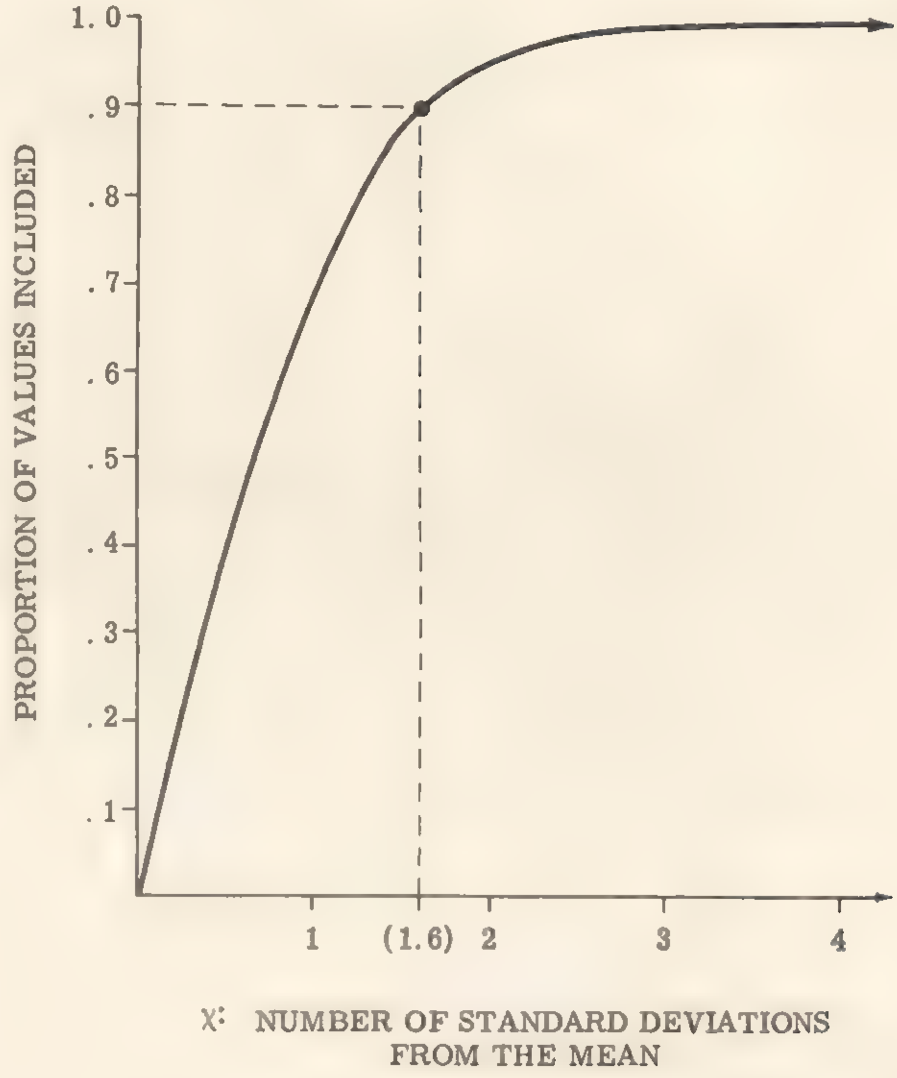
1. There are some things you can say about a collection of data even before you actually collect it. Suppose you were interested, for example, in the heights of students at a particular high school. You could determine the distribution of these heights by measuring and recording the height of each student in the high school. If there were 1, 500 students, your collection of data would consist of __________ observations of a variable called "height." 1, 500

2. Suppose you only knew the heights of ten of the 1, 500 students. Although these ten observations form a collection of data, they could also be considered as part of the larger, complete collection of data. Statisticians use the name **sample** to describe a collection of data which is viewed as part of a larger, complete collection of data. In other words, the collection of __________ heights 10
10/1, 500
would be considered a **sample** in this illustration.

3. Statisticians refer to the **complete** collection of data (of which the sample is a part) as a **population**. Thus, the heights of the ten students would be considered a sample, whereas the heights of the 1, 500 students would be considered a ______________. population

139.

PROPORTION OF VALUES WITHIN χ STANDARD DEVIATIONS OF THE MEAN IN A NORMAL DISTRIBUTION



The dotted lines on the previous graph indicate that

________ of the values in a normal distribution are (.9/ .8) .9

within ________ standard deviations of the mean. (1.6/1.2) 1.6

4. Let's look at another example of the difference between a sample and a population. Suppose you wanted to know which of two candidates for some public office was most preferred by each of the 10, 000 people in your city. If you asked the first 100 people you met on the street to state their preference, you would have a sample consisting of _______ observations from the complete collection of data (population) consisting of __________ observations.

100

10, 000

5. Suppose you were interested in the yearly income of the 200 school teachers in your city. A collection of data consisting of yearly incomes of only 5 of these teachers would be a _______________ (sample/population) if you viewed it as part of the _______________ (sample/population) consisting of the yearly incomes of all 200 teachers.

sample

population

6. It is important to realize that a particular collection of data could be treated as either a sample or population depending upon how you view it. In the previous illustration, for example, the complete collection of data consisting of the yearly income of each of the 200 teachers in your city was viewed as a **population**. Suppose, however, you were interested in the yearly incomes of all the teachers in your state. A **complete** collection of data would now consist of a list of yearly incomes of all the teachers in the state.

This list would now be the population, whereas the list of yearly incomes for the 200 teachers in your city now could be viewed as a _______________ (sample/population) from this population.

sample

137. (Continued)

To find the proportion of values within some number of standard deviations from the mean, simply locate that number along the bottom of the graph. Then find a point on the curve directly above the number you found on the baseline. Finally, find the point on the left side of the graph that is the same height as the point on the curve. This procedure is illustrated by the **dotted lines** on the graph. The vertical dotted line (which goes from the bottom of the graph to the curve) indicates that the dot on the curve is directly **above** the number ______ on 1
1/2
the bottom of the graph.

The horizontal dotted line (which goes from the left side of the graph to the curve) indicates that the dot on the curve is at **the same height** as a proportion of ________, .68
.68/ .78
shown on the side of the graph.

138. The dot on the curve is directly above a particular **number of standard deviations** and at the same height as a **proportion**. This indicates that the proportion of values is within that particular number of standard deviations from the mean of a normal distribution. Thus, the point identified by the dotted lines on the previous graph indicates that the proportion of values within ______ one
standard deviation from the mean of a normal distribution is _______. .68

7. Notice also that the yearly income for any 5 particular teachers in your city ____________ (could/could not) be viewed as a sample from **either** a population consisting of the incomes of all the teachers in your city, or the larger population consisting of the yearly incomes of all the teachers in your state. could

8. A particular collection of data may be viewed as a ____________ (sample/population) if you are comparing that collection to a **larger** collection of data of which it is a part, or as a ____________ (sample/population) if you are comparing it to a **smaller** collection of data which would be included in it. sample population

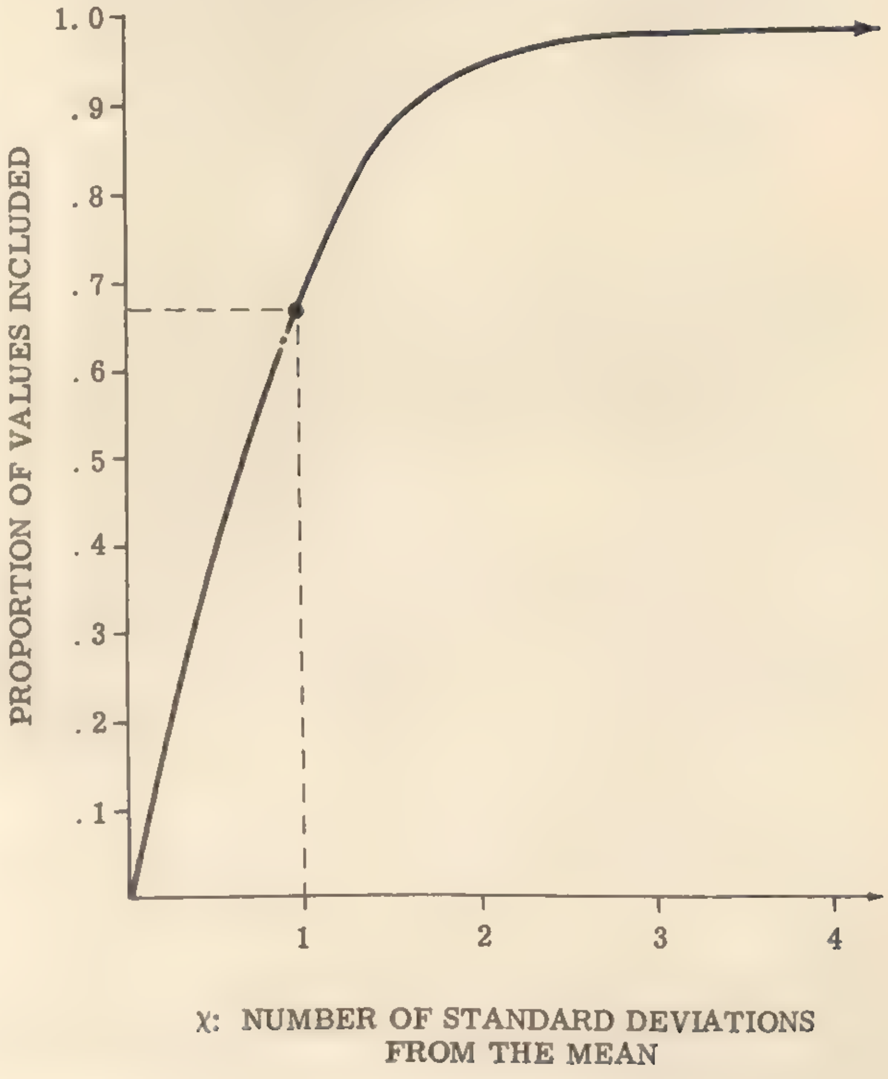
9. To define a population, you could describe what would be included in the complete collection of data. For example, you could describe a collection of data consisting of the ages of all the Democratic presidents. You could describe another collection of data consisting of the ages of all the Republican presidents.

Either one of these collections of data could be considered a population, but they ____________ (would/would not) be the same population(s). would not

10. A particular collection of data may be a sample from more than one population. For example, a list of the ages of all the truck drivers in Detroit could be a s__________ from the population consisting of the ages of all the truck drivers in the United States. The same collection of data could be considered a ____________ from a population consisting of all the male workers in Detroit. sample sample

137. The **theoretical** formula or equation for a normal distribution specifies the proportion of values within any number of standard deviations from the mean. This information is summarized by the following graph.

PROPORTION OF VALUES WITHIN χ STANDARD DEVIATIONS OF THE MEAN IN A NORMAL DISTRIBUTION



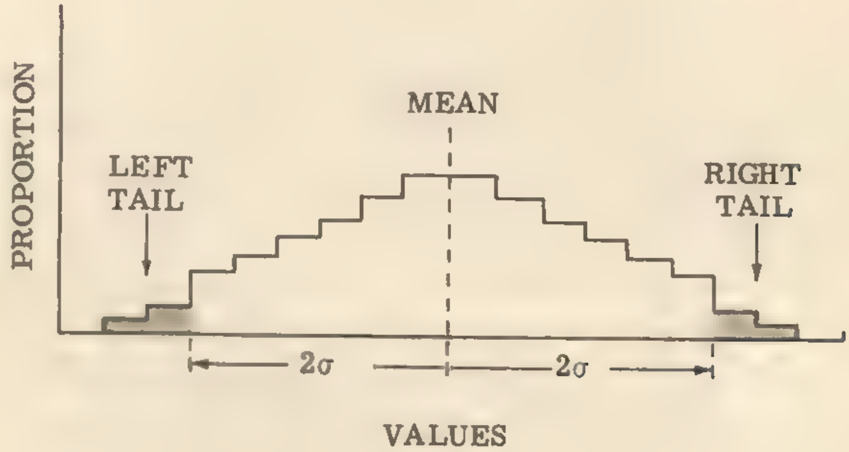
11. The ages of all the female school teachers in Detroit ______________ be a sample from a population would not
would/ would not
consisting of the ages of all the male drivers in the United States. However, the ages of all the female school teachers in Detroit ______________ be a sample would
would/ would not
from a population consisting of the ages of all the people employed in Detroit.

12. Up to this point, we have only considered populations as collections of data that could actually be completely collected. For example, we discussed a population consisting of the heights of all the students in a particular high school. In principle, at least, it ______________ be possible to actually collect a would
would/ would not
complete list of the heights of all the students in the high school.

13. It is sometimes useful to think of a particular collection of data **as if** it were a sample from a larger collection of data consisting of an **unlimited** number of observations. An example of a population consisting of an unlimited number of observations would be a list of the outcomes of an unlimited number of throws of a die (one member of a pair of dice). You could roll a die and record the number of dots showing on the face of the die as the "outcome" of this toss. You could then roll it again, and again, and again, each time recording the number of dots showing on the face of the die. In this way, you could go on producing a list of observed values, without ever specifying where you should stop. A list of outcomes for any specific group of tosses could be viewed as a ______________ from a longer, unlimited list of outcomes, since this smaller collection of data would be part of the larger, **unlimited** collection. sample

136. (Continued)

This is illustrated in the following graph:



The distribution in this graph is **similar**, but not identical, to a normal distribution. We have shaded the columns representing values of 2 standard deviations or more from the mean. Since the distribution is **symmetrical** and since exactly .05 of all the values are 2 σ's or more from the mean, the proportion represented by the shaded area in the left tail is _________ and the proportion in the other tail is _______.

.025

.025

14. If a population consists of a specific number of observations, it is called a **finite** population. The hair color of every person in the United States is a finite population since, in principle at least, you could actually list the hair color or every person in the United States.

Similarly, since there are a limited number of people in the world at any one time, a list of the heights of everyone in the world would consist of a ______ (limited/unlimited) number of observations and would therefore be a **finite** population.

limited

15. A population viewed as consisting of an **unlimited** number of observations is referred to as an **infinite** population. A population consisting of the yearly income of each person in a particular city ______ (would/would not) be an **infinite** population, since the number of people in the city is limited (in principle, you could actually count them all).

would not

16. Suppose you shuffled a deck of playing cards, dealt out five cards, and counted the number of red cards among the five. You could consider this number to be a single observation of a variable whose 6 possible ______ are: 0, 1, 2, 3, 4, and 5. You could shuffle the cards again, deal out five more cards,and count the number of red cards in this new group of five. You could continue this process **indefinitely**. (The procedure for generating this list of observations does not specify any **end** to the list.) In other words, the possible number of observations would be ______ (limited/unlimited). Any specific

values

unlimited

133. The previous table indicates the additional values that are included as you move away from the mean in steps of ________ a standard deviation. 1/2

134. By using the formula for the **normal distribution**, it is possible to construct a table of this sort with "steps" as small as you like. This kind of table is useful because you will find that many of the distributions you encounter are very similar to the theoretical shape of a normal distribution. For example, if you knew a distribution was similar to a normal distribution, you would know that approximately ________ (.68/.95) of all the values were within one standard deviation. .68

135. An important feature of a **normal distribution** is the fact that it is **symmetrical**. Therefore, the .68 of the values within one standard deviation of the mean are equally divided on either side of the mean. In other words, .34 of the values in the normal distribution have a positive deviation of less than one σ, whereas another ________ of the values have a negative deviation less than one σ. .34

136. There are .05 of the values in the distribution which are two or more standard deviations away from the mean. These values are equally distributed in either "tail" of the distribution — .025 in one tail and .025 in the other.

16. (Continued)

group of observations, therefore, could be considered to be a **sample** from a(n) __________ population consisting infinite
finite/ infinite
of the unlimited number of observations.

17. We have seen that every collection of data describes a distribution. The distribution can either be described in absolute terms, by listing the frequency with which each value occurred in the data, or it can be described in relative terms, by listing the p__________ of times each value occurred in the data. proportion

18. A distribution can be described by statistics such as the mean, the median, and the mode, all of which characterize its c__________ t__________; central tendency
or a distribution can be described by statistics such as the range and the variance, which describe its v__________. variability

19. While you may be interested in the distribution of a particular collection of data, for one reason or another you will often have only **part** of that complete collection. In other words, although you are really interested in a __________, you may have only a population
population/ sample
__________. sample
population/ sample

129. The proportion of values within two σ 's of the mean is ________, or ________ percent, since .95 equals 95/100. .95, 95

130. Only ________ of the observed values are 3 standard deviations or more from the mean. (Again, the proportion .01 equals _____ %). .01 1

131. Since .99 of the values are less than 3 σ's from the mean, only 1 - _____, or .01, of the values are 3 or more standard deviations from the mean. .99

132. The **normal distribution** is an example of a **theoretical distribution** since its shape is defined by a mathematical equation and not simply by an actual distribution of data. The equation defining the normal distribution specifies the proportions of values within **any** number of standard deviations from the mean. Thus, the previous table could be expanded to give a more detailed description of a normal distribution.

Number of σ's from the Mean	Proportion of Values
Less than 1/2	.38
Less than 1	.68
Less than 1 1/2	.87
Less than 2	.95
Less than 2 1/2	.98
Less than 3	.99

Notice that each proportion could be converted to a percentage by simply dropping the decimal point, since $.38 = \frac{38}{100} = 38\%$ and $.68 = \frac{68}{100} =$ _____%, and so on. 68

20. Suppose, for example, you were interested in the number of people in the United States with a college education. It would probably be too difficult, too time-consuming, and too expensive to actually list whether or not each person in the United States held a college degree. Suppose, however, you stopped twenty people on a street corner and asked them if they held a college degree. The data you collected would be only **part** of the data in which you were interested. Therefore, the twenty observations would be a ____________ from the ________________ consisting of the education of all the people in the United States.

sample
population

21. Since there is a **specific number** of people in the United States, your list of twenty observations would be a sample from a ____________ population.
finite/infinite

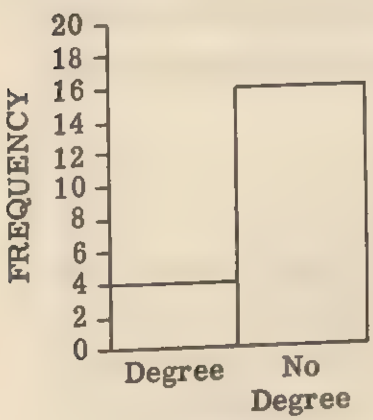
finite

22. Suppose you found that four of the twenty people you had stopped on the street did hold college degrees. You could describe the distribution of your sample by the **absolute** frequency distribution shown in Graph ______,
A/B

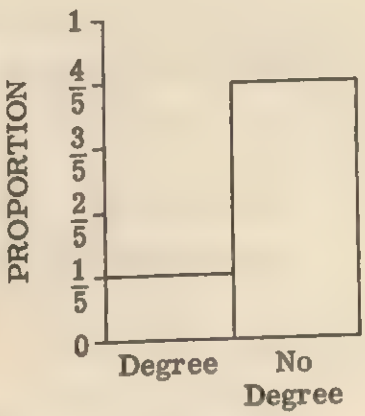
or by the **relative** frequency distribution shown in Graph ______.
A/B

A

B



GRAPH A



GRAPH B

127. (Continued)

This table indicates how many more observations are included as you proceed in 1/2 standard deviation steps away from the mean. Thus, the first row indicates 4/10 of the values were within _______ a standard deviation of the mean. 1/2

The second row indicates ________ - tenths of the 8
observed values were within ______ standard deviation 1
of the mean.

No additional values were included when the limit was extended to include values less than $1\frac{1}{2}$ standard deviations from the mean. However, all of the values are within ______ standard deviations of the mean. 2

128. Statisticians often describe the **shape** of a distribution by indicating the proportion of values within different numbers of standard deviations of the mean. For example, a very important type of distribution, **normal distribution**, can be described in this manner.

Number of Standard Deviations from the Mean	Proportion of Observations
Less than one	.68
Less than two	.95
Less than three	.99

Notice that .68 of all the values in this type of distribution
are within _______ standard deviation of the mean. 1
(The proportions in the table above have been rounded off to the nearest hundreth).

23. If you felt that the sample of 20 people was **typical** or **representative** of the whole population, you might guess that the proportion of people in the United States with college degrees was about ______. $\frac{1}{5}$

$\frac{1}{5}$ / $\frac{4}{5}$

24. In other words, while you don't actually have a complete collection of data, you might feel that the **sample** is **similar** to the ______________. population

25. A guess at some statistic describing the population on the basis of a sample from that population is called a **statistical inference**. You made a ______________ statistical ______________ about the proportion of college inference graduates in the United States on the basis of a sample of 20 people.

26. A **statistic** describing a **population** is often called a **population parameter**. For example, if you had a complete collection of data listing whether or not each person in the United States had a college degree, you could calculate the true proportion of people with college degrees in this population. Since the proportion would be a statistic describing a population, it would be called a population p______________. parameter

27. Any statistic describing a sample is called a **sample statistic**. Thus, the **proportion** of the people with degrees in your **sample** would be a

______________________________. sample statistic

sample statistic/ population parameter

126. One way of **describing** the previous distribution would be to list the **proportion** of values within different distances from the mean. Consider, for example, the following table:

Number of Standard Deviations from the Mean	Proportion of Observed Values
Less than 1	8/10
Less than 2	10/10

The first row indicates that 8/10 of the observed values were closer to the mean than one standard deviation.

The second row indicates that __________ of the observed values were within two standard deviations from the mean. 10/10 (all)

We ______________ (could/could not) have described the collection of ten observations shown in Frame 116 in this manner. could

The data consisted of ten observed values of which only two values were more than one σ away from the mean. Those values were _____ and ______. 0, 20

127. You could describe the same data in more detail if you expanded the previous table as follows:

Number of Standard Deviations from the Mean	Proportion of Observed Values
Less than 1/2	4/10
Less than 1	8/10
Less than 1 1/2	8/10
Less than 2	10/10

28. Suppose you were a shoe manufacturer who had to decide what proportion of his total shoe production should be devoted to each possible shoe size. You would want the distribution of shoe sizes you produced to be as close as possible to the actual distribution of shoe sizes required by the people who could purchase your shoes, since it would not be efficient for you to produce more shoes then you could sell in one size and fewer shoes then could be sold in another size. In principle at least, you could measure the shoe size of every potential customer and thereby determine the actual distribution of shoe sizes in the population. It probably would be out of the question, however, to actually determine the shoe sizes of all these people. Suppose you were able to obtain the shoe sizes of 100 potential customers. You might **guess** something about the distribution of the

_______________ based on the distribution of the — population
sample/population

_______________ you had obtained. — sample
sample/population

29. In other words, on the basis of a sample of 100 shoe sizes from the population, you would be making a statistical in_______________ concerning the distribution of shoe sizes in the population of potential customers. — inference

30. The difference between the largest and the smallest shoe size in the sample of 100 shoe sizes would be the ___________ of the sample. — range

120. X_4 is **below** the mean. Therefore, even though it is the same distance from the mean as is X_2, it has a standard score of ________, instead of 1.82. -1.82

121. Both the observed value of 20 and the observed value of ______ have the same **absolute deviation** from the mean, even though one deviation is positive and the other is negative. 0

122. The only two observed values whose **absolute deviation** from the mean was **greater than** one standard deviation are ______ and ______. 20, 0

123. Since there were only ___________ observed values farther away from the mean than one standard deviation, you would say that ________ of all ten observations had values **less than** one standard deviation from the mean. two

$\frac{2}{10}$ / $\frac{8}{10}$ $\frac{8}{10}$

124. Any standard score **greater** than +1 or **less than** -1 would represent a value more than ________ standard deviation away from the mean. one

125. Out of the ten observed values in the previous table, there were ______ **within** (less than) $\frac{1}{2}$ a standard deviation from the mean. 4

There were _____ values within one σ of the mean. 8

31. The difference between the largest and the smallest shoe size in the population of all potential customers is a statistic that describes the ______________ and would, therefore, be called a p______________ statistic. population population
population/ sample

32. The variance of your **sample** would be a ______________________________. sample statistic
sample statistic/ population parameter

33. If you had a complete collection of data concerning the shoe sizes of all potential customers, you could actually calculate the true variance of the population. This variance would be a population statistic or p__________. parameter

34. Suppose you guessed that the population (parametric) mean was the same as the mean of your sample. You would be using a ______________ statistic as an estimate of a ______________ statistic. sample population
population/ sample

35. You would be making an **inference** about the __________ on the basis of a __________. The difference between your **estimate** of the population mean and the **true** population mean would be your **error of estimate**. population sample

In other words, if the true mean were 10 and your estimate were 9, the **error of estimate** would be 10 - ____, or ____. 9, 1

116. The following table contains the collection of data we considered previously. The fourth column contains the standard scores corresponding to these values.

In other words, the last entry in each row indicates the number of ________ ____________ between the observed value in that row and the mean. In this table, $\bar{x} = 10$ and $\sigma = 5.5$. — standard deviations

OBSERVATION	VALUE	$X - \bar{x}$	STANDARD SCORE
1	15	5	.91
2	20	10	1.82
3	10	0	0
4	0	-10	-1.82
5	15	5	.91
6	10	0	0
7	5	-5	-.91
8	10	0	0
9	5	-5	-.91
10	10	0	0

117. Notice that observation one (X_1) corresponds to a standard score of .91. This means that 15 was slightly ________ (less/more) than one standard deviation from the mean (where $\bar{x} = 10$ and $\sigma = 5.5$). — less

118. Observation two (X_2) corresponds to a standard score of ________, which means that the deviation of X_2 — 1.82
from $\bar{x}$ equals ________ standard deviations. — 1.82

119. X_3 corresponds to a standard score of ______, since the — 0
deviation of observation ______ from the mean was — three
equal to exactly ______ standard deviations. — 0

36. Consider the two collections of data shown below:

Observation	Value
1	10
2	40
3	20
4	25
5	20
6	20
7	40
8	20
9	15
10	25

TABLE A

Observation	Value
1	25
2	10
3	25

TABLE B

Suppose the data in Table A represented a **population** and the data in Table B represented a **sample** from that population. The mode of the population is _____, whereas the mode of the sample is ______.

20
25

37. If you used the sample mode as an estimate of the population mode, your **error of estimate** would equal 20 - _____, or ______.

25, -5

38. If the sample were **typical**, or **representative**, of the population, the difference between a sample statistic and the population statistic would be small. In other words, your e__________ of e______________ would be small if you used this sample statistic as your estimate of the population statistic.

error, estimate

39. However, the sample statistic might be quite different from the population statistic and the error of estimate would then be _____________.
large/small

large

112. Another way of saying that the value 21 is two standard deviations **above** the mean would be to say that the value 21 represents a __________ __________ of ______. standard score, 2

2/-2

113. If the $\overline{x}$ = 10 and σ = 5.5, a **standard score** of -1 would correspond to an observed value of ______. 4.5

114. A formula for the standard score of any distribution can be written as follows:

$$\text{Standard score} = \frac{X - \overline{x}}{\sigma}$$

According to this formula, you simply divide the deviation of the value by the standard deviation. Your answer is the number of standard __________ between the observed value and the mean. deviations

Therefore, the number you obtain by dividing $(X - \overline{x})$ by σ is a __________ score. standard

115. In the example you considered earlier, you saw that the value 21 would represent a **standard score** of 2 in a distribution where $\overline{x}$ = 10 and σ = 5.5.

This is consistant with the formula, since

$$\frac{X - \overline{x}}{\sigma} = \frac{21 - 10}{5.5} = \frac{11}{5.5} = \_\_\_\_\_$$ 2

40. Some samples you obtained could be quite representative of the population because the sample statistic was very similar to the population statistic. However, you could also obtain samples which gave you a poor or distorted picture of the population. It would be useful to know how often different values of the sample statistic would occur if you collected one sample after another. Suppose there were 1000 students at the high school and you knew the height of five particular students. The collection of 1000 heights would be your ________________, whereas population
sample/population
the collection of five heights would be a ________________. sample
sample/population

41. Suppose you had twenty such samples from this population, where each sample was a collection of the heights of five students. If you calculated the mean of each of these samples, the list of these twenty means would be a collection of twenty ________________ sample
sample/population
statistics.

42. The frequencies with which each possible value of the sample mean occurred in this collection of twenty sample statistics define a distribution of sample means. The distribution of these twenty sample means is called a **sampling distribution**, since it is a distribution of sample statistics.

Instead of calculating the mean of each of the twenty samples, you could calculate the variance of each sample. The list of twenty sample variances would also be a collection of twenty ____________ statistics. sample
Therefore, the **distribution** of these twenty sample variances would be a ________________ distribution. sampling

107. The **deviation** of each observed value from the ______ is shown in the third column of the table. — mean, $\bar{x}$

Naturally, according to the definition of the **mean**, the total of all the entries in the third column of the table is ______. — 0

108. The ______ of each of the deviations from the mean is shown in the fourth column of the table. — square

109. To find the ______ of the previous data, you would add all the squared deviations from the mean and divide this total by the number of observations. — variance (σ^2)

110. The sum of all the squared deviations from the mean in the previous data is simply the total of column 4. Thus, $\sum(X - \bar{x})^2$ is equal to ______. — 300

300/400

According to the following formula, the variance of this data is equal to ______ divided by ______. — 300, 10

$$\sigma^2 = \frac{\sum(X - \bar{x})^2}{N}$$

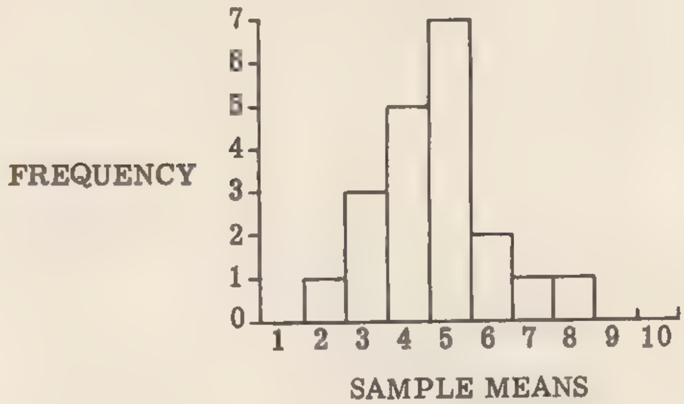
Thus, the variance (σ^2) of the previous data is simply 300 divided by 10, which indicates that the average squared deviation from the mean is ______. — 30

111. The ______ ______ of the previous data is equal to $\sqrt{30}$, or approximately 5.5. — standard deviation

A value of 21 would be exactly ______ standard deviations away from the mean, if σ = 5.5 and $\bar{x}$ = 10. — two

43. Similarly, you could calculate the median, the mode, the range, or the standard error (or any other sample statistic) of these twenty samples. In each case, the distribution of the resulting twenty sample statistics would be called a s____________ distribution. sampling

44. If you collected several samples from some population and calculated the mean of each sample, the following graph might be the frequency distribution of these sample means and would therefore be a ____________ distribution of sample means. sampling



According to this graph, there were ______ samples, 20
10/20

since you calculated a mean for each sample.

45. The largest (**not** the most frequent) sample mean was _____ and the smallest sample mean was ______. 8, 2

46. The most frequently occurring sample mean was ____, 5
since _____ of the twenty samples had that mean. 7
Therefore, the modal value of the sample mean was ___. 5

102. Converting values to **standard scores** is useful when you wish to indicate how many ______________ standard
______________ each value is from the mean. deviations

103. In distributions that are "bell shaped," standard scores of 2 or -2 are __________ (more/less) frequent than are standard less
scores of 1 or -1.

104. Deviations of two standard deviations (standard scores of 2 or -2) are less frequent than deviations of one standard deviation, since most of the values are clustered around the mean in a "_________ __________" bell-shaped
distribution.

105. Suppose you collected the data shown in the following table:

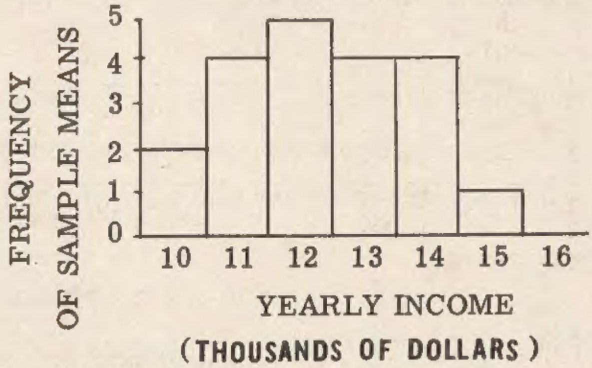
OBSERVATION	VALUE	$X - \bar{x}$	$(X - \bar{x})^2$
1	15	5	25
2	20	10	100
3	10	0	0
4	0	-10	100
5	15	5	25
6	10	0	0
7	5	-5	25
8	10	0	0
9	5	-5	25
10	10	0	0

Since the total of all observed values is ______, you 100
could divide this by ______ to find the mean. Therefore, 10
$\bar{x}$ is equal to ______. 10

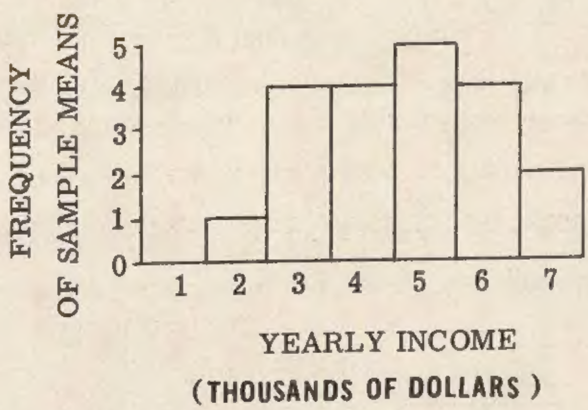
106. The deviation of observation 1 from the mean is
simply $X_1 - \bar{x}$ or, in this case, ____ - _____, which 15, 10
equals ______. 5

47. Suppose you were interested in a population consisting of the yearly incomes of all of the people in a particular city. If you determined the yearly income of ten people you stopped on the street in that city, the resulting list of ten incomes would be considered a ____________ from that population. sample

48. Imagine you collected twenty samples, where each sample consisted of a list of the yearly incomes (to the nearest thousand dollars) of ten people. Furthermore, imagine you collected these twenty samples by stopping people in the lobby of the most expensive hotel in town. Suppose you collected **another** 20 samples on a street corner in the poorest section of town. The graphs shown below could indicate the distribution of sample means in each group of twenty samples.



GRAPH A



GRAPH B

97. A **standard score** is simply the number of standard deviations between a particular value and the mean. If the value is greater than the mean, the standard score is ________ (positive/negative). If a value is less than the mean, the standard score is ________ (positive/negative). — positive / negative

A value identical to the mean has a deviation of 0. Such a value, therefore, is equivalent to a **standard score** of ______. — 0

98. If a value represented a standard score of -2, you would know it was 2 standard deviations ________ (above/below) the mean. — below

99. If $\bar{x} = 10$ and $\sigma = 2$, the value ______ would represent a standard score of 2. — 14

The value 14 represents a **standard score of 2** since it deviates from the mean of 10 by 4. This deviation of 4 equals ________ standard deviations, since $\sigma = 2$. — two

100. The value ______ (6/14) would represent a standard score of -2 in a distribution with $\bar{x} = 10$ and $\sigma = 2$, since ______ minus 10 equals -4, and -4 is equal in absolute size to two standard deviations. — 6 / 6

101. Standard scores of -1 or 1 would **both** represent values which deviated from the mean by one standard deviation. A standard score of ______ (-1/1) would represent a value one standard deviation below the mean, whereas the other standard score would represent a value one standard deviation above the mean. — -1

RESPONSE MASK

Tear this mask off and use it to cover the responses on the right-hand side of each page as you work through this programmed text.

Study the first frame (statement). Then write your response in the space provided or on a separate sheet of paper. NOW move the mask down to reveal the printed (correct) response and check your answer against it. Remember to keep the other responses covered.

Repeat this procedure for each frame. When you have completed a page, slip this mask over the responses on the next page and continue your studies.

AN ADDISON-WESLEY PROGRAMMED TEXT

ADDISON-WESLEY PUBLISHING COMPANY, INC.,
Reading, Massachusetts • Palo Alto • London